

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2002, 17:42:59 ; Search time 34.93 Seconds
(without alignments)
2314.965 Million cell updates/sec

Title: us-09-423-516-2
Perfect score: 4126
Sequence: 1 MWYTKLIPALLQHVLHLL.....VAYYAKWIKHILTYKVPQS 728

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_032802:*

1:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22:	/SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4126	100.0	728	14	AAAR42062
2	4126	100.0	728	17	AAAM00338
3	4126	100.0	728	17	AAAM00340
4	4126	100.0	728	19	AAAW59222
5	4126	100.0	728	19	AAAW58966
6	4126	100.0	728	19	AAAW42998
7	4126	100.0	728	19	AAAM39207
8	4126	100.0	728	14	AAAR39521
9	4119	99.8	728	13	AAAR25676
10	4119	99.8	728	14	AAAR40862
11	4118	99.8	728	13	AAAR20005

12	4116	99.8	728	20	AAW88529	Human hepatocyte g
13	4115	99.7	727	12	AAAR10656	Hepatic parenchyma
14	4113	99.7	728	20	AAW88532	Human hepatocyte g
15	4112	99.7	728	12	AAAR15623	Human leukocyte-de
16	4112	99.7	728	20	AAW88531	Human hepatocyte g
17	4111	99.6	728	20	AAW88530	Human hepatocyte g
18	4110	99.6	728	17	AAAR87525	Mutant hepatocyte
19	4109	99.6	728	17	AAAR87523	Mutant hepatocyte
20	4108	99.6	728	14	AAAR40863	Competative inhibi
21	4107	99.6	728	14	AAAR14243	Human hepatocyte g
22	4106	99.5	728	17	AAAR87524	Mutant hepatocyte
23	4096	99.3	728	15	AAAR47727	Hepatocyte growth
24	4095	99.2	728	17	AAAR87522	Mutant hepatocyte
25	4088.5	99.1	723	11	AAAR07144	Tumour cytotoxic f
26	4088.5	99.1	723	12	AAAR15624	Human leukocyte-de
27	4088.5	99.1	723	15	AAAR57026	Human wild-type tu
28	4088.5	99.1	723	16	AAAR82685	Tumour cytotoxic f
29	4088.5	99.1	723	13	AAAR76690	Human plasmalogen-
30	4085.5	99.0	723	13	AAAR21142	Human TCF-II. Hom
31	4085.5	99.0	723	15	AAAR57027	Human modified tum
32	4085.5	99.0	723	15	AAAR57028	Human modified tum
33	4081.5	98.9	723	13	AAAR25677	Recombinant human
34	4078.5	98.8	723	19	AAAR59923	Human leukocyte-de
35	4073	98.7	728	13	AAAR25160	Human HGF. Homo s
36	4066	98.5	728	13	AAAR12792	Human hepatocyte g
37	4066	98.5	728	19	AAAW48221	Human recombinant
38	4066	98.5	728	20	AAAW97041	A human hepatocyte
39	4066	98.5	728	22	AAAG68090	Human hepatocyte g
40	4066	98.5	728	22	AAAB84517	Amilo acid sequenc
41	4064.5	98.5	723	13	AAAR29819	TCF-II. AAR29819
42	4064.5	98.5	723	17	AAAR99688	TCF mutant having
43	4064.5	98.5	723	17	AAAR99689	TCF mutant having
44	4061	98.4	728	13	AAAR21976	Human Hepatocyte g
45	4061	98.4	728	14	AAAR52942	Human hepatocyte g

ALIGNMENTS

RESULT	1
AAAR42062	standard; Protein; 728 AA.
XX	
AC	AAAR42062;
XX	
DT	15-NOV-1993 (first entry)
XX	
DE	Vascular endothelial cell growth enhancer protein.
XX	
KW	Enhance; growth; vascular endothelial cell; human; tumour; cell line;
KW	HUCCA-II; HUCCA-II; blood vessel; wounds; burns; decubitus;
KW	post-operative tissue damage; drug; cardiac angiodopathy.
XX	
OS	Homo sapiens.
XX	
PN	EP550296-A.
XX	
PD	07-JUL-1993.
XX	
PF	27-NOV-1992; 92EP-0403199.
XX	
PR	28-NOV-1991; 91JP-0337999.
XX	
PA	(TERU) TERUMO CORP.
XX	
PI	Adachi M, Harada K, Hirahara I, Sudo T;
XX	
DR	WPI: 1993-215669/27.
XX	
DR	N-PSDB; AAQ45702.
XX	
PT	Vascular endothelial cell growth factor protein - used for
PT	promoting angiogenesis in the treatment of cardiac angiodopathy,
PT	wounds, burn injuries, postoperative tissue damage etc.

XX Claim 6; Page 18-21; 44pp; English.

CC This sequence represents a single chain protein which selectively
 CC enhances the growth of vascular endothelial cells. This protein
 CC was produced by the human tumour cell line HUCCA-II or HUCCA-III.
 CC This protein enhances the formation of new blood vessels and may be
 CC used to enhance healing of wounds, burns injuries, decubitus or
 CC post-operative tissue damage. It may also be used as a drug for
 CC cardiac angiopathy.

XX
 XX
 SQ Sequence 728 AA;

Query Match 100.0%; Score 4126; DB 14; Length 728;
 Best Local Similarity 100.0%; Pred. No. 9.5e-277;
 Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTVTLPLALLQHVLLHLLELPALPYAEGQRKRNTIHEFKSAKTLIKIDPALKIK 60
 1 mmtvklpalllqhvllhllelpalpyaegqrkrntihefkksaktlikidpalkik 60
 61 TKKVNADQCANRCTRNKGLPFTCKAFVBDKARKQCLMPFNSMSGVKKERGFEDLYE 120
 61 tkkvnadqcanrcrnkglpftckafvbdkarkqclwfpfnsmsgvkketgfedfelye 120
 QY 121 NKDYIRNCIIGKRSYKGYSTIRKSGIKQOPWSSMIPHEHSFLPSSYRGDLOENYCRNP 180
 121 nkdyirnciigkrsykygystirksgikcqpwsmlphensflpsyrqgdldenyccrnp 180
 QY 181 RGEEGPWCFTSNPEVRYEVDIPQCEVECMTCNGESYRGMLDHTESGKICQRMDHQTP 240
 181 rgeegpwcftsnpevryevcdipqcevecmtcngesyrglmhdhtesgkicqrmdhqtp 240
 QY 181 RGEEGPWCFTSNPEVRYEVDIPQCEVECMTCNGESYRGMLDHTESGKICQRMDHQTP 240
 181 rgeegpwcftsnpevryevcdipqcevecmtcngesyrglmhdhtesgkicqrmdhqtp 240
 QY 241 HRHAKLPERYDPKGGDDNYCRNPQOPRPMCTLDPHRTWMECAIKTCADNMNDTVL 300
 241 hrhaklperrydpkggddnycrnpdqpprmctldphrtwmeckaitcandmndtv 300
 QY 241 hrhklperrydpkgfdhnycrnpdqpprmctldphrtwmeckaitcandmndtv 300
 301 ETECTIQGEGRYGTWNTINNGICQWRDSQYPHEHMTPEFKCKDLRENYCRNPDS 360
 301 etectiqgegrygtwntlnngicqwrdsyphemtpenfckdlrenyccrnpds 360
 QY 301 etectiqgegrygtwntlnngicqwrdsyphemtpenfckdlrenyccrnpds 360
 301 etectiqgegrygtwntlnngicqwrdsyphemtpenfckdlrenyccrnpds 360
 QY 361 ESPWCFTTDPNIRVGYCSQIPNCDMSHGDCYRGNGKNYMGNLSTGRSLTCSMDKME 420
 361 espwcttcdpnirvgycsqipncdmshgdcyrgngknymgnlstgrsltcsmdkme 420
 QY 361 espwcttcdpnirvgycsqipncdmshgdcyrgngknymgnlstgrsltcsmdkme 420
 361 espwcttcdpnirvgycsqipncdmshgdcyrgngknymgnlstgrsltcsmdkme 420
 QY 421 DLHRIIFEPDASKINENYCRNPDDAHGPMCTGNPLIPMPYCPISRCESGTPPIVNL 480
 421 dlhriifepdasklnenycrnpddahgpmctgnplipmpyccpiscrgstppivnl 480
 QY 421 dlhriifepdasklnenycrnpddahgpmctgnplipmpyccpiscrgstppivnl 480
 421 dlhriifepdasklnenycrnpddahgpmctgnplipmpyccpiscrgstppivnl 480
 QY 481 DHPVISCATKQOLRVVNGIPRTNIGMWVSLRYRNKHICGSLIKESVWLTAROCPPSRD 540
 481 dhpviscatkqlrvvngiprtlnigmwslryrnkhicgslikesvwltaarocpsrd 540
 QY 481 dhpviscatkqlrvvngiprtlnigmwslryrnkhicgslikesvwltaarocpsrd 540
 481 dhpviscatkqlrvvngiprtlnigmwslryrnkhicgslikesvwltaarocpsrd 540
 QY 541 LKDYPAWLGIHDVHGRDECKQVNLVNSQVYGPBGSDLVLMKLARPAVLDFVSTIDLP 600
 541 lkdyawlgihdvhgrdekckqvnlvnsqvygpbgsdlvmlklarparvldfvstidlp 600
 QY 541 lkdyawlgihdvhgrdekckqvnlvnsqvygpbgsdlvmlklarparvldfvstidlp 600
 541 lkdyawlgihdvhgrdekckqvnlvnsqvygpbgsdlvmlklarparvldfvstidlp 600
 QY 601 NYGCTIPKTSYSGVGMGTGLINVDGILRVAHLTYMGNEKXSOHHRGVTINSEIICG 660
 601 nygctipktsysgvgtglinydgilrvahltymgnekxsohhrgvtnseicg 660
 QY 601 nygctipktsysgvgtglinydgilrvahltymgnekxsohhrgvtnseicg 660
 601 nygctipktsysgvgtglinydgilrvahltymgnekxsohhrgvtnseicg 660
 QY 661 AEKISSGCEGDYGPPLVCEQHKRMVNLGVIVPGRGCAIPNPGIIFVRAVYAKVHKKII 720
 661 aekissgcegdypplvceqhkrmvnlgvivpgrgcalpnpgiifvrayakvhwkii 720
 QY 661 aekissgcegdypplvceqhkrmvnlgvivpgrgcalpnpgiifvrayakvhwkii 720
 661 aekissgcegdypplvceqhkrmvnlgvivpgrgcalpnpgiifvrayakvhwkii 720
 QY 721 LTYKVPQS 728
 721 ltykvps 728
 QY 721 ltykvps 728
 721 ltykvps 728

RESULT 2
 AAM00338

ID AAM00338 standard; protein; 728 AA.

XX AC AAM00338;

XX 27-JUN-1997 (first entry)

DE Human hepatic parenchymal cell growth factor.

XX Human; hepatic; parenchymal cell; growth factor; HGF; treatment;

KW prevention; ischemia; ischaemic; disease; reperfusion; disorder;

KW blood; liver; transplant; acute; failure; ischemia; ischemic;

KW organ.

XX Homo sapiens.

OS WO9632960-A1.

XX 24-OCT-1996.

PD 19-APR-1996; 96WO-JP01065.

XX 21-APR-1995; 95JP-0096994.

XX (MITU) MITSUBISHI CHEM CORP.

PI Gemba M, Yonehana T;

DR WPI; 1996-48555/48.

XX Hepatic parenchymal cell growth factors - used as remedy and/or

PT preventive medicine for ischaemic diseases

PS Claim 9; Pages 7-10; 16pp; Japanese.

XX The present sequence is the human hepatic parenchymal cell growth

CC factor (HGF), which has a molecular weight of 76-92 kD estimated by

CC SDS-PAGE, parenchymal cell growth activity, which is lost following

CC heat treatment at 80 degrees C for 10 minutes and strong affinity

CC towards heparin. It may be used to treat and prevent ischaemic

CC diseases, e.g. blood reperfusion disorder, ischaemic reperfusion

CC disorders in liver transplant, acute liver failure and organ

CC disorder during organ transplant. For adults, it is administered at

CC a level of 1 microg to 10 mg/kg/day, preferably 10-1000

CC microg/kg/day.

CC A pig liver derived cultured epithelial cell line was cultured in

CC DMEM, cow embryo serum and N-2-hydroxyethylpiperazine ethane

CC sulphonic acid to confluence. The medium was then changed to serum

CC and glucose free DMEM, and the oxygen concentration reduced to less

CC than 2%. The cells were cultured for 6 hours, and then for another

CC hour at 95% air/5% carbon dioxide (reoxygenation). HGF suppressed free

CC lactic acid dehydrogenase (LDH) by 49.6% when added before

CC hypoxiation, and by 60.9% when added before reoxygenation, at 50

CC ng/ml (free LDL was 100% in the absence of HGF).

XX Sequence 728 AA;

Query Match 100.0%; Score 4126; DB 17; Length 728;

Best Local Similarity 100.0%; Pred. No. 9.5e-277;

Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTVTLPLALLQHVLLHLLELPALPYAEGQRKRNTIHEFKSAKTLIKIDPALKIK 60
 1 mmtvklpalllqhvllhllelpalpyaegqrkrntihefkksaktlikidpalkik 60
 QY 61 TKKVNADQCANRCTRNKGLPFTCKAFVBDKARKQCLMPFNSMSGVKKERGFEDLYE 120
 61 tkkvnadqcanrcrnkglpftckafvbdkarkqclwfpfnsmsgvkketgfedfelye 120
 QY 61 tkkvnadqcanrcrnkglpftckafvbdkarkqclwfpfnsmsgvkketgfedfelye 120
 61 tkkvnadqcanrcrnkglpftckafvbdkarkqclwfpfnsmsgvkketgfedfelye 120
 QY 121 NKDYIRNCIIGKRSYKGYSTIRKSGIKQOPWSSMIPHEHSFLPSSYRGDLOENYCRNP 180
 121 nkdyirnciigkrsykygystirksgikcqpwsmlphensflpsyrqgdldenyccrnp 180
 QY 121 nkdyirnciigkrsykygystirksgikcqpwsmlphensflpsyrqgdldenyccrnp 180
 121 nkdyirnciigkrsykygystirksgikcqpwsmlphensflpsyrqgdldenyccrnp 180

```

OY 181 RGEGRFPWCFSTNPEVRYEVDIPQCESEVEMCTNGESYRGIMDHTESGKICQRMHQTP 240
    |||||||
Db 181 rgeegrpwcfstnpevryevcdipqceveomctngesyrgimhtesgkicqrmhqtp 240
OY 241 HRIKFLPERYPDKGFDNDYCRNPDGQPRPWCYTLDPHTFMRYCAIKTCADNTMNDTVPL 300
    |||||||
Db 241 hrnkflperypdkgfdndycrnpdgprpwcyltldphtfmycaiktcadntmndtvp 300
OY 301 ETEECIOGEGRGYRGTNTIWNIGIPCORMDQYRPHENDMTPEMKCKDLNENYCRNDGS 360
    |||||||
Db 301 etteciogggyrgygtvntiwnigipcorwdsqyphendmtpemckckdlenycrndgs 360
OY 361 ESPWCFETDNRIRGYCSQIPNCDMSHGQDCYRGNKNGMGNLSQTFSSGILTCSMWDMKME 420
    |||||||
Db 361 espwcfetdnrirygysqipncdmsingdcyrgnknymgnlsqtfsgiltsmwdkme 420
OY 421 DLHRIHFWERDASKLNNENYCRNPDGAGFPWCYTGNPLRPMDCPISRCGDTTPITVNL 480
    |||||||
Db 421 dlhrhifwerdasklneenycrnpddahgfpwcytgnpplrmdypisrcgdttpitvnl 480
OY 481 DHPYISCAKTKQLRVVNGIFPTNIGMWSLRYNKNIKCGSLIKESWVLTARCCFERSD 540
    |||||||
Db 481 dhpviscaktkqlrvvngifptnigmwslrynkhiqsgslikeswvltarccfersd 540
OY 541 LKDYEAMLGIDYHGRGDEKCOVLNYSQLYGPEGSDLYLMLKARPAVLDDFVSTIDLP 600
    |||||||
Db 541 lkdyeamlgidvhgrgdekcyvlmysqlyvgpessdlylmlkarpavlddfvstldlp 600
OY 601 NYGCTIPEKTSQSVYGMGYGLINYDGLRVAHLIYMGNEKCSQHRGKVTYLNSEICAG 660
    |||||||
Db 601 nygctipektsesvgygylinydglirvahlyimgnekcsqhrgkvtlineseicag 660
OY 661 AERTIGSPGCGDYGGLVCSQHKRMVLYIVPGRGALINRRPFIPIRVAYAAVHIKII 720
    |||||||
Db 661 aertigspcgdygglyvcsqhkrmvlyivpgrcalinrrpfiirvayaaikhkii 720
OY 721 LTYKVPQS 728
    |||||||
Db 721 ltykvpqs 728

RESULT 3
ID AAM00340 standard; protein; 728 AA.
XX AAM00340:
XX AAM00340:
XX 09-DEC-1996 (first entry)
XX Wild type hepatocyte growth factor.
XX Human: hepatocyte growth factor; HGF; huHGF; serum; proteolytic cleavage;
XX pro-hormone; beta subunit; alpha subunit; kringle domain; prothrombin;
XX plasminogen; catalytic domain; serine protease; HGF variant;
XX HGF receptor; malignancy; chronic HGF receptor activation.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FH Cleavage-site 494..495
XX FT /label= "Proteolytic cleavage-site"
XX FT /note= "generates alpha and beta subunits"
XX FT Disulfide-bond 487..604
XX FT Peptide 1..54
XX FT /note= "prepro-sequence"
XX FT Peptide 1..31
XX FT /note= "Hydrophobic signal peptide"
XX FT Domain 128..206
XX FT /label= kringle_1_domain
XX FT Domain 211..288
XX FT /label= kringle_2_domain
XX FT Domain 303..383

```

```

FT FT /label= kringle_3_domain
FT Domain 391..464
FT /label= kringle_4_domain
FT Modified-site 294
FT Modified-site /label= N-1-linked-glycosylation_site
FT Modified-site 402
FT Modified-site /label= N-1-linked-glycosylation_site
FT Modified-site 566
FT Modified-site /label= N-1-linked-glycosylation_site
FT Modified-site 653
FT /label= N-1-linked-glycosylation_site
FT Misc-difference 494
FT /note= "Position of opt. substitution, esp. Glu, Asp or Asn"
FT Misc-difference 495
FT /note= "Position of opt. substitution, esp. Tyr or Phe"
FT Misc-difference 534
FT /note= "Position of opt. substitution"
FT Misc-difference 673
FT /note= "Position of opt. substitution"
FT Misc-difference 692
FT /note= "Position of opt. substitution"
XX XX US5547856-A.
XX PD 20-AUG-1996.
XX PF 18-MAY-1992; 92US-0884811.
XX PR 13-JUL-1993; 93US-0087783.
XX PR 18-MAY-1992; 92US-0884811.
XX PR 18-MAY-1992; 92US-0885971.
XX PA (GETH ) GENEWTECH INC.
XX XX Godowski PJ, Lokker NA, Mark MR;
XX PI WPI; 1996-392634/39.
XX DR New hepatocyte growth factor variants - are resistant to in vivo
XX PT proteolytic cleavage into a 2-chain form, useful as HGF antagonists
XX PS Disclosure; Column 51-56; 39pp; English.
XX XX This sequence represents the wild type full length sequence of human
XX CC hepatocyte growth factor (HGF). HGF is isolated from human serum
XX CC and is a disulphide linked heterodimer derived by proteolytic cleavage
XX CC of this pro-hormone form between residues 494 and 495. This generates a
XX CC molecule composed of an alpha-subunit of 440 amino acids (mol. wt. 69
XX CC kD) and a beta-subunit of 234 amino acids (mol. wt. 34 kD). The alpha
XX CC and beta subunits are encoded by a single open reading frame. The alpha
XX CC subunit contains four kringle domains based on their homology to
XX CC kringle-like domains in other proteins, e.g. prothrombin, plasminogen,
XX CC The beta subunit shows high homology to the catalytic domain of serine
XX CC proteases. However two of the three residues which form the catalytic
XX CC triad of serine proteases are not conserved in HGF. Therefore, the
XX CC precise function of the beta chain remains unknown. The invention
XX CC includes HGF variants which retain HGF receptor binding activity without
XX CC having the biological activity of wild-type HGF. They can be used for
XX CC the treatment of pathological conditions associated with the activation
XX CC of a HGF receptor such as malignancies associated with chronic HGF
XX CC receptor activation.
XX SQ Sequence 728 AA;
OY 1 MWVTKLIPALLQVHLHLPLAIPAEQGRKRRNTHEFFKSATTKIKIPAIKIK 60
Db 1 mvvtklilpallqvlhllplaiapyaeqgrkrntthfeksaktklikipalkik 60
Query Match 100.0%; Score 4126; DB 17; Length 728;
Best Local Similarity 100.0%; Pred. No. 9, 5e-277;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 61 TTKVNTADQCANRCTRNNKGLPFTCKAFVFDKARKOCLMFPFNSMSGYKKEFGHEPDLYE 120
DB 61 Ttkvntadqcanctrnnkglpftckafvfdkarkqclwprfnsmssgykkefghefdlye 120
QY 121 NKDYIRNCIIGKGRSYKGTVSITKSGIKCQPWSSMI PHEHSFLPSSYRGKDLQENYCRNP 180
DB 121 nkdyirnciigkgrsykgtvsitksigikcqpwsmslphesflpsysrgkdlqenyccrnp 180
QY 181 RGEEGSPWCFTSNPEVREVCIDIPQCESEVCMTCNGESYRGIMDHTEESKTCQRMDHOTP 240
DB 181 rgeegspwcftsnpevrevcidipqcevecmctcngesyrgimdhthesktcqrmdhotp 240
QY 241 HRHKFLPERYPDKGFDDNYCRNPDGQPRPWCYTLLDPHTRMETCAIKTCADNTMNDVPL 300
DB 241 hrhkflperypdkgfdndyccrnpdgqprpwcyclldphtmetcaiktcadntmndvpl 300
QY 301 ETTETICOGGEGYRGTVNTIMNGICQRMDSOY PHEHDMTPENFKCKDLRENYCRNPDGS 360
DB 301 etteciggggegrygtvntlmngicqrmdsyphedmtpefkckdlrenyccrnpdgs 360
QY 361 ESPWCFTTDPNIRVGYCSQIPNCDMSHGDCYRGNGKNYMGNLSTQTRGILCSMDKNME 420
DB 361 espwcttdpnirvgycsqipncdmsbgdcyrgngknymgnlstqtrgilcsmdkme 420
QY 421 DLHRHIFWEPDASKLNNENYCRNPDDAHGPCYTGNPLIPMDYCPISRCGDTPTPIVNL 480
DB 421 dlhrhifwepdasklneenycrnpddahgpcytnplipmdycpisrcgdtptpivnl 480
QY 481 DHPYISCKTQOLRVNVCIPRTNIGMNVSLRYRKAHICGSLKESVLTAROCFSPSRD 540
DB 481 dhpysiscktkqlrvnvciprtnigmnvslryrkahicgslkesvltarcfspd 540
QY 541 LKDYEAWLGIHDVHGRGDEKCKQVNLNSQLYEGEGSDLVLMKLARPAVLDPFSTIDLP 600
DB 541 lkdyeawlgihdvhgrgdekckqylnvnsqlyyeggsdlvmlklarpavldpfstidlp 600
QY 601 NYGCTIPEKTSQSYGMYGTLINYDGLLRVAHL YIMGNKCSOHHRGKVTLLNESELCAG 660
DB 601 nygctipektsqsygygtlinydglrrvahlyimgnkcsqhhrgkvtllneselcag 660
QY 661 AEKIGSGCEDEDYGGPLVCEBQHKMRWLVGIVPRGCAIPRRGIFRVAAYAKIHKIT 720
DB 661 aekigsgcededyggplvcebhkmrvlgvprgcaiprrgifrvaayakihkii 720
QY 721 LTRYKPOS 728
DB 721 ltrykpos 728

```

```

PA (NAKA/) NAKAMURA T.
XX Asami O, Hagiya M, Ihara I, Nakamura T, Sakauchi M;
PI Seki T, Shimizu S, Shimonishi M;
XX WPI: 1998-429650/37.
DR N-PSDB; AAV53626.
XX
PT Preparation of recombinant hepatocyte growth factor polypeptide - by
PT culturing mammalian cells transformed with vector containing human
PT leukocyte-derived HGF gene
PS
PS Claim 1; Fig 2a-c; 30pp; English.
XX
CC This is the amino acid sequence for human leukocyte-derived
CC hepatocyte growth factor (HGF) HLC3, deduced from a cDNA clone
CC (see AAV53626) obtained from a leukocyte cDNA library. Another
CC leukocyte-derived HGF, HLC2 (see AAV59233), was identified that
CC differs from HLC3 by having amino acid residues 162-166 deleted.
CC A claimed method of producing a HGF comprises: transforming
CC mammalian cells (preferably CHO cells) with a recombinant
CC expression vector that has a promoter (preferably the SV40 early
CC promoter), a gene coding for human leukocyte-derived HGF, and the
CC dihydrofolate reductase gene; (b) culturing the transformed cells
CC in the presence of successively elevated concentrations of
CC methotrexate; and (c) recovering the HGF polypeptide from the
CC culture supernatant. HGF polypeptides may be used as hepatocyte
CC cultivation reagents, liver regeneration promoters, in basic
CC research on liver function, research on the action of various
CC hormones and drugs on hepatocytes, research on the carcinogenesis
CC mechanism of hepatoma, clinical diagnostic reagents using an
CC antibody against the polypeptide and therapeutic drugs for liver
CC disease.
XX
SQ Sequence 728 AA;

```

```

Query Match 100.0%; Score 4126; DB 19; Length 728;
Best Local Similarity 100.0%; Pred. No. 9, 5e-277;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MWVTKLLPALLLQVHLHLHLPLAIPYABGQRKRRTIHEFKSAKATLLIKIDPALKIR 60
DB 1 mwvtkllpalllqvhllhlhlplaiipyabgqrkrntihefksaktllikidpalkir 60
QY 61 TTKVNTADQCANRCTRNNKGLPFTCKAFVFDKARKOCLMFPFNSMSGYKKEFGHEPDLYE 120
DB 61 Ttkvntadqcanctrnnkglpftckafvfdkarkqclwprfnsmssgykkefghefdlye 120
QY 121 NKDYIRNCIIGKGRSYKGTVSITKSGIKCQPWSSMI PHEHSFLPSSYRGKDLQENYCRNP 180
DB 121 nkdyirnciigkgrsykgtvsitksigikcqpwsmslphesflpsysrgkdlqenyccrnp 180
QY 181 RGEEGSPWCFTSNPEVREVCIDIPQCESEVCMTCNGESYRGIMDHTEESKTCQRMDHOTP 240
DB 181 rgeegspwcftsnpevrevcidipqcevecmctcngesyrgimdhthesktcqrmdhotp 240
QY 241 HRHKFLPERYPDKGFDDNYCRNPDGQPRPWCYTLLDPHTRMETCAIKTCADNTMNDVPL 300
DB 241 hrhkflperypdkgfdndyccrnpdgqprpwcyclldphtmetcaiktcadntmndvpl 300
QY 301 ETTETICOGGEGYRGTVNTIMNGICQRMDSOY PHEHDMTPENFKCKDLRENYCRNPDGS 360
DB 301 etteciggggegrygtvntlmngicqrmdsyphedmtpefkckdlrenyccrnpdgs 360
QY 361 ESPWCFTTDPNIRVGYCSQIPNCDMSHGDCYRGNGKNYMGNLSTQTRGILCSMDKNME 420
DB 361 espwcttdpnirvgycsqipncdmsbgdcyrgngknymgnlstqtrgilcsmdkme 420
QY 421 DLHRHIFWEPDASKLNNENYCRNPDDAHGPCYTGNPLIPMDYCPISRCGDTPTPIVNL 480
DB 421 dlhrhifwepdasklneenycrnpddahgpcytnplipmdycpisrcgdtptpivnl 480

```


QY 481 DHPVISC AKTKOLRVNNGIPTRTNIGMMVSLRYRNKHHICGSLIKESWVLTAROCFSPSRD 540
 |||||||
 Db 481 dhpvisacktkqlrvnvngiprtinigmwslryrnkhhicgslikeswvltarqcfpsrd 540
 QY 541 LKDYEAMLGIDHVGKDECKOVLNVSQLYVGPESDYLMLKARPAVLDDEVSTIDLP 600
 |||||||
 Db 541 lkdyeawlgidhvgkdeckqvlvnsqlyvgpesdylmlkarpauiddfvstidlp 600
 QY 601 NYGCTIPEKTSVYGGVGYGLINVOGLIRVAHLXYINGNEKCSOHHGKVTLNSEIICAG 660
 |||||||
 Db 601 nycctipektsvgygyglinydglirvahlyimgnekcsqnhgkvclnseicag 660
 QY 661 AEKIGSGPCGEGDYGRLVCEQHKKRMVLYIVPGRCALPNRPGIEFVAVAYAKMIHKII 720
 |||||||
 Db 661 aekigsgpcegdylgprlvceqhkrmvlyivpgrcalpnrpglfvravayakwihkii 720
 QY 721 LTYKVPQS 728
 |||||||
 Db 721 ltykvpqs 728

RESULT 5

AAW58696
 ID AAW58696 standard; protein; 728 AA.

XX AAW58696;
 AC

XX 08-SEP-1998 (first entry)
 DT

XX Human hepatocyte growth factor.
 DE

XX Human; hepatocyte growth factor; HGF; high blood sugar.
 KM

XX Homo sapiens.
 OS

XX JP10167981-A.
 PN

XX 23-JUN-1998.
 PD

XX 09-DEC-1996; 96JP-0328357.
 PF

XX 09-DEC-1996; 96JP-0328357.
 PR

XX (MITU) MITSUBISHI CHEM CORP.
 PA

XX WPI; 1998-408607/35.
 DR

New preventive and/or treating agent - comprises hepatocyte growth factor, used for diseases caused by high blood sugar

XX Claim 5; Page 4-6; 6pp; Japanese.
 PS

XX The present sequence represents human hepatocyte growth factor (HGF).
 CC

XX A preventive and/or treating agent for the diseases caused by high
 CC blood sugar has been developed. The agent comprises HGF as an active
 CC component. The preventive and/or treating agent causes no excessively
 CC low blood sugar.

XX Sequence 728 AA;
 SQ

Query Match 100.0%; Score 4126; DB 19; Length 728;
 Best Local Similarity 100.0%; Pred. No. 9.5e-277;

Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWYTKLLPALLGHVLLHLLLPALPAIYAGQRRKRNTHIEFKKSAKTTLLIKIDPALKIK 60
 |||||||
 Db 1 mwvtkllpalllghvllhlllpalpaiaegqrkrntihfkkksaktllikidpalkik 60
 QY 61 TKKVNTADQCANRCTRRKKGIPFTCKAFVFDKARKQCIAMPFNSSSGVKKKEFGHEFDLYE 120
 |||||||
 Db 61 tkkvntadqcanrctrnkgipftckafvfdkarqciawfnsmsgvkkefghetdlye 120

QY 121 NKDYIRNCIIGKGRKYGVSTIKSGIKQFWSMIPHEHSFLPSSYRGKDIDQENYCRNP 180
 |||||||
 Db 121 nkdyirnciigkgrkygvstitsgikqfwsmi p h e h s f l p s s y r g k d i d q e n y c r n p 180
 QY 181 RGEEGGFWCFTSNPEVRYEVCDDIPQCSBECVTCNGESYKGLMDHTSGKICQRMHQTP 240
 |||||||
 Db 181 rgeeggfwcftsnpevryevcdipqcsbecvtcngesyrglmdhtsgkicqrmhqt p 240
 QY 241 HRHKEFLERYPDDKFDNNYCNBNPDGPRPWCYTLDPHTRMRYCAIKKCANDTMNDTPVL 300
 |||||||
 Db 241 hrhkeflerypdkfdnnycnbnpdgprpwcycl d p h t r m r y c a i k k c a n d t m n d t p v l 300
 QY 301 ETPECIOGCGEYGTAVNTIWNIGI PCORWDSQY PHEHDMTPENKCKDLRENYCRNPDGS 360
 |||||||
 Db 301 ettecio g g e y g t v n t i w n i g i p c o r w d s q y p h e h d m t p e n k c k d l r e n y c r n p d g s 360
 QY 361 ESPWCFITTDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNSLQTSRGLTCSKWDK NME 420
 |||||||
 Db 361 espwcfitt d p n i r v g y c s q i p n c d m s h g o d c y r g n g k n y m g n s l q t s r g l t c s m w d k n m e 420
 QY 421 DLHRHIFMEPDSKLTNNYCNBNPDGDAHGPWCYGNPLIPWDYCPISRCEGDTPTIVNL 480
 |||||||
 Db 421 dlhrhifmepdsklt n n y c n b n p d g d a h g p w c y g n p l i p w d y c p i s r c e g d t p t i v n l 480
 QY 481 DHPVISC AKTKOLRVNNGIPTRTNIGMMVSLRYRNKHHICGSLIKESWVLTAROCFSPSRD 540
 |||||||
 Db 481 dhpvisacktkqlrvnvngiprtinigmwslryrnkhhicgslikeswvltarqcfpsrd 540
 QY 541 LKDYEAMLGIDHVGKDECKOVLNVSQLYVGPESDYLMLKARPAVLDDEVSTIDLP 600
 |||||||
 Db 541 lkdyeawlgidhvgkdeckqvlvnsqlyvgpesdylmlkarpauiddfvstidlp 600
 QY 601 NYGCTIPEKTSVYGGVGYGLINVOGLIRVAHLXYINGNEKCSOHHGKVTLNSEIICAG 660
 |||||||
 Db 601 nycctipektsvgygyglinydglirvahlyimgnekcsqnhgkvclnseicag 660
 QY 661 AEKIGSGPCGEGDYGRLVCEQHKKRMVLYIVPGRCALPNRPGIEFVAVAYAKMIHKII 720
 |||||||
 Db 661 aekigsgpcegdylgprlvceqhkrmvlyivpgrcalpnrpglfvravayakwihkii 720
 QY 721 LTYKVPQS 728
 |||||||
 Db 721 ltykvpqs 728

RESULT 6

AAW42998
 ID AAW42998 standard; protein; 728 AA.

XX AAW42998;
 AC

XX 23-JUL-1998 (first entry)
 DT

XX Recombinant human hepatocyte growth factor (HGF).
 DE

XX hepatocyte growth factor; HGF; W/O/W emulsion; blood half life;
 KM therapeutic; drug; hepatic tissue; hepatic disease; acute hepatitis;
 KW chronic hepatitis.
 KW

XX Homo sapiens.
 OS

XX JP10007587-A.
 PN

XX 13-JAN-1998.
 PD

XX 24-JUN-1996; 96JP-0163063.
 PF

XX 24-JUN-1996; 96JP-0163063.
 PR

XX (AZUM/) AZUMA H.
 PA

XX (MITU) MITSUBISHI CHEM CORP.
 PA

XX (MIYA-) MIYAZAKI KEN.
 PA

XX (TSUB/) TSUBOUCHI H.
 PA

XX WPI: 1998-126093/12.
 XX Preparation of emulsified composition - contains hepatocyte growth
 PT factor
 XX
 XX Claim 5: Pages 7-8; 9pp; Japanese.
 XX
 CC The present sequence represents a recombinant human hepatocyte growth
 CC factor (HGF). The protein is used in the inner aqueous phase of a
 CC W/O/W emulsion. The emulsion also contains an oil phase and an outer
 CC aqueous phase. The emulsified composition can considerably extend the
 CC blood half life of HGF. The emulsion is used in therapeutical drugs to
 CC deliver HGF to hepatic tissues in patients. It is useful for the
 CC treatment of hepatic diseases such as acute hepatitis and chronic
 CC hepatitis.
 CC
 XX
 XX Sequence 728 AA:

Query Match 100.0%; Score 4126; DB 19; Length 728;
 Best Local Similarity 100.0%; Pred. No. 9.5e-277;
 Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTWKLLPALLOHVLHLHLPLAIPYAEGRKRRMTIHEFKSAKTTLIKIDPALKIK 60
 DB 1 mwtckllpalllqhvllhlllplaiyaegqrkrntlnhefkksaktllikidpalkik 60
 QY 61 TKKVNADQCANRCTRNKGLPFTCKAFVFDKARKQCLMPENSSSGYKKEGHEFDLYE 120
 DB 61 tkkvnadqcanrctrnkgjlpftckafvfdkarqclwfpnsssgykkefghfdlye 120
 QY 121 NKDYIRNCITIGKRSYKGVTSITSGIKCOPWSSMIPHEHSEFLSSYRGKDLQENYCNRP 180
 DB 121 nkdyirncilgkrsygvtsitsgikcqpssmiphehsflpsysrgkdlqenycnrlp 180
 QY 181 RGEEGPWCFTSNPEVREYVCDIPQCESEVECMTCNGESYRGIMDHTESGKICQRMHQTP 240
 DB 181 rgeegpwcftsnpevreyevcdipqcevecmtcngesyrgimhtesgkicqrmwbqtp 240
 QY 241 HRHKFLPERYDCKGFDNDYCNRPDQPRPWCYTLDPHTRWECALKTCADMTMDTVPL 300
 DB 241 hrhkfleperrydckgfdndycnrdpqrprwcytlphtrewecalktcaadmtdtdvpl 300
 QY 301 ETTTCIOGCGRGYGTNTWNGIIPCQRWDSQYRPHENDMPENFKCDLRENYCRNPDGS 360
 DB 301 ettciogcgrgygtntwngiipcqrwdsqyrphendmpenfkcldlrenycrnpdgs 360
 QY 361 ESPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYRNGKNYMGNLQSRTSGLTCSMWDMKME 420
 DB 361 espwctftdpnirvgycsqipncdmsbgdcyrgngknymgnlsqtrsgltcsmwdkme 420
 QY 421 DLHNRITWEPASKLNEYCNRPDDANGPWCYTGNPILIPMDYCRISCEGDTPTIYNL 480
 DB 421 dlhnrityepasklneycnrpdngpwcytgnpilipmdycriscegdtptrtynl 480
 QY 481 DHPYISCAKTKQLRVNNGIPRTNIGMMVSLRYRNKNIICGSLIKESWYLAROCFSPSRD 540
 DB 481 dhpyscaktkqlrvnngiprtnigmmvslryrnknicgslikeswylarocfspd 540
 QY 541 LKDYEAMLGIDHVGREGDECKQVNLNSQVLYGREGSDLVLMKLARPAVLDPVSTIDL 600
 DB 541 lkdyeamlgidvhgredckqylnlsqvlygregsdlvmlklarpvldpfvstidl 600
 QY 601 NYGCTITEKTCGSYYGMWYTLINWDGLLRVAHLITMGENKCSQHNRKATVLNNESEIAG 660
 DB 601 nygctitektsyvgmwytlinwdglrrvahlitmgenkcsqhnrkavtlnneeseia 660
 QY 661 AEKIGSGCEGDYCGPLVCEQHKMMVNLGVIVPGRCALIPNRPGLFVVAUYAKIHKKII 720
 DB 661 aekigsgcedygcplvceqhkmmvnlgvivpgrcalipnrpglfvvaayakihkii 720
 QY 721 LTYKVPQS 728

DB 721 ltykvpqs 728

RESULT 7

AAW39207

ID AAW39207 standard; protein; 728 AA.

AC AAW39207;

DT 11-MAY-1998 (first entry)

DE Human hepatocyte growth factor protein.

KW Hepatocyte growth factor; HGF; parenchymatous cell proliferation factor;

KW insulin-like; human; preventative agent; therapeutic agent; glycogen;

XX hypoglycaemia.

XX Homo sapiens.

PN JP10007586-A.

PD 13-JAN-1998.

PF 05-JUN-1996; 96JP-0142839.

PR 26-APR-1996; 96JP-0108263.

PA (MITU) MITSUBISHI CHEM CORP.

DR WPI: 1998-133629/13.

PT Composition for treating conditions caused by insulin agonists -

PS e.g. hypoglycemia and glycogen diseases

XX Claim 5; Page 5-6; 8pp; Japanese.

XX This sequence represents a human hepatocyte growth factor (HGF), also

XX known as a hepat parenchymatous cell proliferation factor which can be

XX used as a preventative and/or therapeutic agent for diseases caused by

XX insulin-like action, e.g. hypoglycemia and glycogen disease. This factor

XX has activity to proliferate hepat parenchymatous cells and has a strong

XX affinity with heparin.

XX Sequence 728 AA:

Query Match 100.0%; Score 4126; DB 19; Length 728;
 Best Local Similarity 100.0%; Pred. No. 9.5e-277;
 Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMTWKLLPALLOHVLHLHLPLAIPYAEGRKRRMTIHEFKSAKTTLIKIDPALKIK 60
 DB 1 mwtckllpalllqhvllhlllplaiyaegqrkrntlnhefkksaktllikidpalkik 60
 QY 61 TKKVNADQCANRCTRNKGLPFTCKAFVFDKARKQCLMPENSSSGYKKEGHEFDLYE 120
 DB 61 tkkvnadqcanrctrnkgjlpftckafvfdkarqclwfpnsssgykkefghfdlye 120
 QY 121 NKDYIRNCITIGKRSYKGVTSITSGIKCOPWSSMIPHEHSEFLSSYRGKDLQENYCNRP 180
 DB 121 nkdyirncilgkrsygvtsitsgikcqpssmiphehsflpsysrgkdlqenycnrlp 180
 QY 181 RGEEGPWCFTSNPEVREYVCDIPQCESEVECMTCNGESYRGIMDHTESGKICQRMHQTP 240
 DB 181 rgeegpwcftsnpevreyevcdipqcevecmtcngesyrgimhtesgkicqrmwbqtp 240
 QY 241 HRHKFLPERYDCKGFDNDYCNRPDQPRPWCYTLDPHTRWECALKTCADMTMDTVPL 300
 DB 241 hrhkfleperrydckgfdndycnrdpqrprwcytlphtrewecalktcaadmtdtdvpl 300
 QY 301 ETTTCIOGCGRGYGTNTWNGIIPCQRWDSQYRPHENDMPENFKCDLRENYCRNPDGS 360
 DB 301 ettciogcgrgygtntwngiipcqrwdsqyrphendmpenfkcldlrenycrnpdgs 360

Db 301 ettecigqggygrtvtntlmwngipcqrdwsqyphndtptenfckckdlrenycrnpgds 360
 QY 361 ESPWCFTTDPNIRVYCSQIIPNCDMSHGDCYRGNGKNYMGNTLSQTRSGLTCSMDKNME 420
 Db 361 espwctftdpnirvycsqipncdmshgdcyrngknymgnlsgtrsgltcsmwcknme 420
 QY 421 DLHRHIFWEPDASKLLENYCRNPDDDAHGPWCYTGNPLIPWDYCPISRCGDTPTIVNL 480
 Db 421 dlhrhifwepdasklleenycrnppddahgpwcytgnplipwdycpisrcegdtptrlvnl 480
 QY 481 DHPVISCATKQLRVVNGIPTRNIGMWVSLRRKNKHICGSLIKESWVLTARQCPSRD 540
 Db 481 dhpviscatkqlrvvngipttrnigmwslrrknkhicgslikeswvltarqcpisrd 540
 QY 541 LKDYEAWLGIHDYHGRDECKQVNLVNSQLVYGPESGDLVLMKLARPAVLDDFTVSTDLR 600
 Db 541 lkdyeawlgihdyhgrdeckqvlvnsqlyvypesgdlvmlklarpavlddftvstcldr 600
 QY 601 NYGCTIPEKTSQSVYGMGYTGLINYGDLRLVAHLIYMGNEKCSQHRHGKVTLINESEICAG 660
 Db 601 nygctipektsqsvygytglyngdlrlvahlyimgneksqhrhgkvtlineselcag 660
 QY 661 AAKIGSGPCEGDYGGPLVCEQHKRMVLYIVPGRGCAIPNRGIFVRAVYAKWIKHII 720
 Db 661 aekigsqcegdyygplvceqhkrmvlyivpgrgcaipnrpgifvrayayakwihkii 720
 QY 721 LTYKVPQS 728
 Db 721 ltykvpqs 728

RESULT 8

AAR39521
 ID AAR39521 standard; Protein; 728 AA.

XX AC AAR39521:
 XX DT 20-SEP-1993 (first entry)
 XX DE Hepatocyte growth factor.
 KW HGF, side effect; cancer; tumour; chemotherapy; radiotherapy;
 KW carcinostatic.
 OS Synthetic.
 XX WO9308821-A.
 XX 13-MAY-1993.
 PF 05-NOV-1992; 92WO-JP01433.
 XX 07-NOV-1991; 91JP-0321412.
 PA (NAKA/) NAKAMURA T.
 PI Nakamura T;
 DR WPI; 1993-167384/20.
 DR N-PSDB; AAQ46040.
 PT Side effect inhibitor for cancer therapy - containing hepatocyte
 PT growth factor and inhibiting side effects resulting from
 PT treatment by chemotherapy, radiotherapy
 PS Disclosure; Fig 1; 36pp; Japanese.
 CC Since HGF alleviates damage to normal cells and tissues, it can
 CC inhibit side effects in treating cancers by chemotherapy,
 CC radiotherapy, etc. With it, it is possible to conduct more intense
 CC cancer therapy and to improve the carcinostatic effects. It is
 CC extremely useful in the field of clinical medicine.

SQ Sequence 728 AA:

Query Match 99.9%; Score 4120; DB 14; Length 728;
 Best Local Similarity 99.9%; Pred. No. 2,5e-276;
 Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MWYTKLIPALLLQHVLLHLLPLAIPYAEGRKRRRTIHEFKSAKTLIKIDPALIK 60
 Db 1 mwytklipalllqhvllhllplaiptyaeqrkrrntihfeksaktlikidpalkik 60
 QY 61 TKRVNTADOCANBCTRNKGLPTCKAFVDFDKARKKOCILFPFNSMSGVKKFGEHEDLYE 120
 Db 61 tkrvntadocanbctrnkglptckafvdfdkarkkocilfpmssgvykkfgehfedlye 120
 QY 121 NKDYIRNCITIGKRSYKGYTITKSGIKCQPNSSMTIPHEHSFLPSSYRGCDLQENCRNP 180
 Db 121 nkdyirncitigkrsygytitsgikcqpssmiphehsflpsyrghdldqenycrnp 180
 QY 181 RGEEGPWCFTSNPEVRYEVCDIPOCSEVECMTCNGESYRGILMDHTESGKICQRMHQTP 240
 Db 181 rgeegpwcftsnpevryevcdipocsevecmtcngesyrgilmdhtesgkicqrmdhqtpr 240
 QY 241 HRHKFLPERYPDKGFDNNYCRNPDGQRPWCYTLDPTHTMEYCAIKTCADNTMNDIVPL 300
 Db 241 hrnkflperypdkgfdnnycrnpdgqrpwcytldphtmeycailtccadntmndivpl 300
 QY 301 EETECTOGGEGYRGTYNTIWNGLPCQRMDSQYRPHEDMTPEAFKCKDLRENYCRNPDGS 360
 Db 301 ettectoggegyrgtyntiwnglpcqrdwsqyphndmtpeafkckdlrenycrnpgds 360
 QY 361 ESPWCFTTDPNIRVYCSQIIPNCDMSHGDCYRGNGKNYMGNTLSQTRSGLTCSMDKNME 420
 Db 361 espwctftdpnirvycsqipncdmshgdcyrngknymgnlsgtrsgltcsmwcknme 420
 QY 421 DLHRHIFWEPDASKLLENYCRNPDDDAHGPWCYTGNPLIPWDYCPISRCGDTPTIVNL 480
 Db 421 dlhrhifwepdasklleenycrnppddahgpwcytgnplipwdycpisrcegdtptrlvnl 480
 QY 481 DHPVISCATKQLRVVNGIPTRNIGMWVSLRRKNKHICGSLIKESWVLTARQCPSRD 540
 Db 481 dhpviscatkqlrvvngipttrnigmwslrrknkhicgslikeswvltarqcpisrd 540
 QY 541 LKDYEAWLGIHDYHGRDECKQVNLVNSQLVYGPESGDLVLMKLARPAVLDDFTVSTDLR 600
 Db 541 lkdyeawlgihdyhgrdeckqvlvnsqlyvypesgdlvmlklarpavlddftvstcldr 600
 QY 601 NYGCTIPEKTSQSVYGMGYTGLINYGDLRLVAHLIYMGNEKCSQHRHGKVTLINESEICAG 660
 Db 601 nygctipektsqsvygytglyngdlrlvahlyimgneksqhrhgkvtlineselcag 660
 QY 661 AAKIGSGPCEGDYGGPLVCEQHKRMVLYIVPGRGCAIPNRGIFVRAVYAKWIKHII 720
 Db 661 aekigsqcegdyygplvceqhkrmvlyivpgrgcaipnrpgifvrayayakwihkii 720
 QY 721 LTYKVPQS 728
 Db 721 ltykvpqs 728

RESULT 9

AAR25676
 ID AAR25676 standard; protein; 728 AA.

XX AC AAR25676:
 XX DT 20-JAN-1993 (first entry)
 XX DE Recombinant human hepatocyte growth factor.
 KW HGF; enhance growth; preparing transgenic animals; hepatic disease;
 KW clinical diagnostic reagent; drug.


```

Db      61 tkkvtadgancrctrnkqjpfctckafvdkarqclwfpnsmssgykkefgnefdlye 120
QY      121 NKQYIRNCIIIGKGRSYKGTYSITKSGIKQCPWSSMIPIHEHSFLPSSYRGKDLQENYCRNP 180
Db      121 nkdyirnciigkgrsykgtvtsitksqikcpwssmiphehsflpsyrqkdlqenycrnp 180
QY      181 RGEEGSPWCFSTSNPEVREYCDIPQCESEVECMTCNGESYRGLMDHTSGKICQKQWMDQTP 240
Db      181 rgeegspwcfstsnpevreyevcdipqcevecmctcngesyrglmdhtesgkicqkwdqtp 240
QY      241 HRHKFLPERYPDKGFDDNYCRNPDGPRPWCYTLDPRTRMEYCAIKTCADNTMNDTPVL 300
Db      241 hrhkflperypdkgfdndycrnpdgprpwcyltdprtrmeycaiktcadntmndtvp 300
QY      301 ETTECIGOGEGYRGYNTWTNNGIPCORWDSQYRPHENDMPENKCKDLRENTYCRNPDGS 360
Db      301 ettecigqgegrygtvntwtngipcqrwdsqyrphendmpenkckdlrenycrnpdgs 360
QY      361 ESPWCFITTDPIRIRVGYCSQIPNCDMSHGDCYRGNGKNYMGNTLSQTSRGLTCSMWKDKME 420
Db      361 espwcfittdpiirirvgysqipncdmshgdcyrgngknymgntlsqtsrsgltcsmwdkme 420
QY      421 DLHRHIFWEPDASKLNNENYCRNPDDAHGPWCYTGNPLIPWDYCPISRCEDTTPPTIVNL 480
Db      421 dlhrhifwepdasklneenycrnpddahgpwcytgnplipwdycpisorcedtptl 480
QY      481 DHPVISCARTKQLRVNNGIPIRTNIGMWNVSLRYRNKHKICGSLIKESWVLTAROCPSRD 540
Db      481 dhpviscartkqlrvnngipirtnigwmvslryrnkhicgslikeswvltaarqcfpsrd 540
QY      541 LKDYEAWLGIHDVHGRGDEKCOVLNLSQLYVGEPSDLVLMKLARAVLDDPFTSIDLP 600
Db      541 lkdyeawlgihdvhgrgdekcoylvlnlsqlyvgepsdlvlmklaravlddftvsidlp 600
QY      601 NYGCTIPEKTSQSYVGYGTYGLINVDGLLRVAHLIYMGNEKCSQHHGKVTLLNSEICAG 660
Db      601 nygctipektsqsyvgygtyglinydgllrvahllymgnekcsqhgvkvlenseicag 660
QY      661 AEKIGSPCEGDYGGPLVCEQHKRMVLYGVLPGRGCAIPNRPGLFVRVAYAAVIAHKII 720
Db      661 aekigspcegdyygplvceqhkrmvlygvlpgrgcaipnrglfrvayayakwihkii 720
QY      721 LTKYKVPQS 728
Db      721 ltkykvps 728

```

```

XX      1 MWVTKLIPALLQVHLHLPLTAIPYAEQRRKRRNTIHEKKSATTLTKIDPALKIK 60
DR      N-PSDB; AAQ20049.
XX      PT Agent conty. hepatocyte growth factor and carrier - used for
XX      PT treating renal diseases and promoting nephrocyte growth and as
XX      PT diagnostic for renal diseases
PS      Disclosure; Fig 1; 15pp; English.
XX      CC Human HGF comprises an alpha-chain of 440 amino acids and a
XX      CC beta-chain of 234 amino acids. There are 4 kringle domains in the
XX      CC alpha-chain, similar to that of plasmin; the beta-chain has about
XX      CC 37 per cent homology with the beta-chain of plasmin having serine
XX      CC protease activity. Homology of the amino acid sequence of rat HGF
XX      CC and human HGF is 91.6 per cent in the alpha-chain and 88.9 per cent
XX      CC in the beta-chain. HGF has been found to be an agent for nephrocyte
XX      CC growth and is useful as a treatment for renal failure.
SQ      Sequence 728 AA:
Query Match 99.8%; Score 418; DB 13; Length 728;
Best Local Similarity 99.9%; Pred. No. 3,4e-276;
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 MWVTKLIPALLQVHLHLPLTAIPYAEQRRKRRNTIHEKKSATTLTKIDPALKIK 60
Db      1 mwvtklilpallqvlhlpltpaiyaegqrkrntlhekksatllkikdpalkik 60
QY      61 JKRYNTADQCANRCTRNKGLPFTCKAFVDFKARKQCLMPFPNSMSGYKKEFGHEFDLYE 120
Db      61 tkkvtadgancrctrnkqjpfctckafvdkarqclwfpnsmssgykkefgnefdlye 120
QY      121 NKQYIRNCIIIGKGRSYKGTYSITKSGIKQCPWSSMIPIHEHSFLPSSYRGKDLQENYCRNP 180
Db      121 nkdyirnciigkgrsykgtvtsitksqikcpwssmiphehsflpsyrqkdlqenycrnp 180
QY      181 RGEEGSPWCFSTSNPEVREYCDIPQCESEVECMTCNGESYRGLMDHTSGKICQKQWMDQTP 240
Db      181 rgeegspwcfstsnpevreyevcdipqcevecmctcngesyrglmdhtesgkicqkwdqtp 240
QY      241 HRHKFLPERYPDKGFDDNYCRNPDGPRPWCYTLDPRTRMEYCAIKTCADNTMNDTPVL 300
Db      241 hrhkflperypdkgfdndycrnpdgprpwcyltdprtrmeycaiktcadntmndtvp 300
QY      301 ETTECIGOGEGYRGYNTWTNNGIPCORWDSQYRPHENDMPENKCKDLRENTYCRNPDGS 360
Db      301 ettecigqgegrygtvntwtngipcqrwdsqyrphendmpenkckdlrenycrnpdgs 360
QY      361 ESPWCFITTDPIRIRVGYCSQIPNCDMSHGDCYRGNGKNYMGNTLSQTSRGLTCSMWKDKME 420
Db      361 espwcfittdpiirirvgysqipncdmshgdcyrgngknymgntlsqtsrsgltcsmwdkme 420
QY      421 DLHRHIFWEPDASKLNNENYCRNPDDAHGPWCYTGNPLIPWDYCPISRCEDTTPPTIVNL 480
Db      421 dlhrhifwepdasklneenycrnpddahgpwcytgnplipwdycpisorcedtptl 480
QY      481 DHPVISCARTKQLRVNNGIPIRTNIGMWNVSLRYRNKHKICGSLIKESWVLTAROCPSRD 540
Db      481 dhpviscartkqlrvnngipirtnigwmvslryrnkhicgslikeswvltaarqcfpsrd 540
QY      541 LKDYEAWLGIHDVHGRGDEKCOVLNLSQLYVGEPSDLVLMKLARAVLDDPFTSIDLP 600
Db      541 lkdyeawlgihdvhgrgdekcoylvlnlsqlyvgepsdlvlmklaravlddftvsidlp 600
QY      601 NYGCTIPEKTSQSYVGYGTYGLINVDGLLRVAHLIYMGNEKCSQHHGKVTLLNSEICAG 660
Db      601 nygctipektsqsyvgygtyglinydgllrvahllymgnekcsqhgvkvlenseicag 660
QY      661 AEKIGSPCEGDYGGPLVCEQHKRMVLYGVLPGRGCAIPNRPGLFVRVAYAAVIAHKII 720
Db      661 aekigspcegdyygplvceqhkrmvlygvlpgrgcaipnrglfrvayayakwihkii 720

```

OY 721 LTYKVPQS 728
|||||
Db 721 ltykvpqs 728

RESULT 12

AAW8529

ID AAW8529 standard; protein: 728 AA.

AAW8529;

02-MAR-1999 (first entry)

Human hepatocyte growth factor (HGF) variant 1.

Human; HGF; hepatocyte growth factor; variant; HGF receptor; medicine; heparan sulphate proteoglycan; cancer.

Homo sapiens.
Synthetic.

Key Location/Qualifiers

Misc-difference 73 /label= R73E /note= "wild-type Arg is replaced with Glu"

Misc-difference 76 /label= R76E /note= "wild-type Arg is replaced with Glu"

W09B51798-A1.

19-NOV-1998.

07-MAY-1998: 98WO-GB01318.

10-MAY-1997: 97GB-0009453.

(IMCR) IMPERIAL CANCER RES TECHNOLOGY.

Birchmeier W, Cherard E, Hartmann G;

WPI. 1999-024213/02.

Hepatocyte growth factor variants - binding to hepatocyte growth factor receptor, useful in medicine e.g. for cancer therapy

Claim 6; Page -: 75pp; English.

The present sequence represents a human hepatocyte growth factor (HGF) variant. The HGF variants of the invention bind to the HGF receptor, but are substantially unable to bind heparan sulphate proteoglycan. The HGF variants are useful in medicine e.g. variants having the same effects on target cells as wt-HGF can be administered to patients requiring HGF, while antagonistic variants can be administered to patients requiring HGF antagonists. They are especially useful for treating cancer. Some of the variants have a longer circulatory half-life in vivo and a greater mitogenic activity than wt-HGF in rats. They are therefore useful therapeutically as wt-HGF, but may have superior effects in vivo, e.g. give greater tissue penetration and ability to reach cell or tissue compartments.

Note: This sequence is not provided in the specification; it has been created by modifying the wild-type HGF sequence.

Sequence 728 AA:

Query Match

Best Local Similarity 99.8%; Score 4116; DB 20; Length 728;

Matches 726; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MWYTKLLPALLLQHVLLHLPLIAIPYAEQGRKRRNTHERKSKYTLIIDLAKTK 60
|||||

Db 1 mwylkllpalllqhvllhlllllpiaipyaeqgrkrrntlhefkkaektlkldpalkk 60
OY 61 TKKVNADOCANRCRNNKGLPFTCAFYVDKARKOCLWEPFNSMSGVKKKEGHEFDYE 120
Db 61 tkkvnadocanrcrnnkglpftcakyvdarkoclwepfnsmsgvkkkeghefdye 120
OY 121 NKDYIRNCIIGKRSYKGVSTKSGIKQPPWSSMIPIHEHSLPSRYGKLDGENYCRNP 180
Db 121 nkdyirnciigkrsykgvstksigikppwssmiipiheshlpsrygkldgenycrnp 180
OY 181 RSEEGPWCFTSNPEVRYEVCIDIPQSEVECTNGESYRGLMHTESGKLCOSRMDQTP 240
Db 181 rseegpwcftsnpevryevcidipqsevectngesyrglmdhtesgklcosrmdqtp 240
OY 241 HRRKFLPERYPKGFDDNYCRNPQSPRPKCYTLDPIHMEYCAIKTCADMTNDTDPVL 300
Db 241 hrkflperypkghfddnycrnpqsprrpkcytldpihmyeacaiktcadmtndtdvpl 300
OY 301 ETEECIGQEGYRGTVNTINNGIIPQRMDSQYPHEHDMTPENFKCKDLRENYCRNPDS 360
Db 301 eteeci gqeg yrgtvntinngiipqrm dsqyphehdmtpenfkcckdlrenycrnpds 360
OY 361 ESPMCFITDPIRYGCSQIPNCDMSHGQCYRNGKNYWNLSQTRSGLTCSMDPKNME 420
Db 361 espmcfitdpirygcsqipncdms hgcycyrngknynwnlsqtrsgltcsmdpk nme 420
OY 421 DLRRHIFWEPDASKLNNYCRNPDDDAHGPWCYTGNIPLIPWDYCPISRCGDTPTIVNL 480
Db 421 dlrrhifwepdaskl nnycrnpdddahgpwc ytgniplipwdycpisrcgdtptivnl 480
OY 481 DHPYISCAKTKOLRYVNGIPIRTNIGMWSLRKRNKHCIGSLLIKESWLTARQCFPSRD 540
Db 481 dhpyscaaktkolryvngipirtnigmwslrk rnkhcigsllikeswltarqcfpsrd 540
OY 541 LKDYEAWLGIHDYHGRDEKCOVLNWSOLVYGPESDYLIMKLAPAYLDQFVSTIDLP 600
Db 541 lkdyeawlgihdyhgrdekcovl nwsolv ygpesdylimklapayldqfstidlp 600
OY 601 NYGCTIPEKTSYGVGWTGTLINYGDLRAVHLXIMGNEKSOHHKRGVLTNSESICAG 660
Db 601 nygctipektsygvwgwtgtl inygd lra vhlximgneksohhkrgvlt nsesicag 660
OY 661 AEKIGSGPEGDYGCPLYCEOHKRMVGLVIVPGCAIPNRPGLFVAYAYAKYIKITII 720
Db 661 aekigsgpegdygcp lyceohkrmv glvivpgcaipnrpglfvayayak yikiti 720
OY 721 LTYKVPQS 728
Db 721 ltykvpqs 728

RESULT 13

AAR10656

ID AAR10656 standard; protein: 727 AA.

AAR10656;

16-APR-1991 (first entry)

Hepatic parenchymal cell growth factor.

Hepatic parenchymal cell growth factor; cirrhosis.

Homo sapiens.

Key Location/Qualifiers

Peptide 30..727 /label= hGHF active fragment

Peptide 32..727 /label= hGHF active fragment

FT EP412557-A.
PN

PD 13-FEB-1991.
 XX 10-AUG-1990: 90EP-0115397.
 XX 27-JUL-1990: 90JP-0200898.
 PR 11-AUG-1989: 89JP-0209449.
 PR 03-APR-1990: 90JP-0088592.
 XX (MITU) MITSUBISHI KASEI CORP.
 PA Kitamura N, Miyazawa K, Datkubara Y, Tsoubouchi H, Naka D;
 PI Takahashi, Matsui R, Yoshiyama Y;
 XX WPI: 1991-045716/07.
 DR N-PSDB; AAQ10489.
 PT Hepatic parenchymal cell growth factor - produced by recombinant
 PT methods, useful for e.g. treating cirrhosis
 PT Disclosure; fig 2; 46pp; English.

CC This hepatic growth factor and its active fragments (see feature
 CC table) are encoded by gene(s) contained within the expression plasmid,
 CC PKCRGF-2. This is used to transform host cells for recombinant
 CC expression of the human hepatic parenchymal cell growth factor (hHGF).
 CC hHGF is useful in the treatment of hepatic diseases, e.g. cirr-
 CC hosis. It is produced in large amounts, stably using this method.
 CC
 XX Sequence 727 AA:
 SQ

Query Match 99.7%; Score 4115; DB 12; Length 727;
 Best Local Similarity 99.9%; Pred. No. 5.5e-276;
 Matches 726; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 WTKLPLALLQHYLHLHLPLAIPYAEQGRKRRNTIEFKKSAKTYLTKIDPALKIKT 61
 DB 1 WVKLPLALLQHYLHLHLPLAIPYAEQGRKRRNTIEFKKSAKTYLTKIDPALKIKT 60
 QY 62 KKVNTAQQCANRCRRNNGLPFTCAFPEDKARAKCCLMPFMSMSGKKEGHEFDLYEN 121
 DB 61 KKVNTAQQCANRCRRNNGLPFTCAFPEDKARAKCCLMPFMSMSGKKEGHEFDLYEN 120
 QY 122 KDYIRNCIIGKRSYKSTVSTTKSGIKQPPWSMIPHSHPSSYRGKDLQENYCRNPR 181
 DB 121 KDYIRNCIIGKRSYKSTVSTTKSGIKQPPWSMIPHSHPSSYRGKDLQENYCRNPR 180
 QY 182 GEEGPMCFSTSNPEYREYVCDIPQSEVECMTCNGESYRGIMDHTESGKICQRMDDHOTPH 241
 DB 181 GEEGPMCFSTSNPEYREYVCDIPQSEVECMTCNGESYRGIMDHTESGKICQRMDDHOTPH 240
 QY 242 RHKFLPERYEDKDDNYCRNPDQPRPWCYTLDPHTRMEYCAIKTCADNMTDTPYLE 301
 DB 241 RHKFLPERYEDKDDNYCRNPDQPRPWCYTLDPHTRMEYCAIKTCADNMTDTPYLE 300
 QY 302 TTETIOGEGYRGCTVNTIWMGICQRMDSOYPREHMTPEMFCKDLRENYCRNPDSE 361
 DB 301 TTETIOGEGYRGCTVNTIWMGICQRMDSOYPREHMTPEMFCKDLRENYCRNPDSE 360
 QY 362 SPWCFETDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYGMNLSQTRSGLTCSMMDKNMED 421
 DB 361 SPWCFETDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYGMNLSQTRSGLTCSMMDKNMED 420
 QY 422 LHRITFEPDASKINENYCRNPDADAGPWCYTGNPILPWYCPISRCGEGTPTIYVLD 481
 DB 421 LHRITFEPDASKINENYCRNPDADAGPWCYTGNPILPWYCPISRCGEGTPTIYVLD 480
 QY 482 HPVISCACKTOLRVVNGIPRTFNIGMMWVSLRYRNKHIGGSLIKESWVLTAROCFPPSD 541
 DB 481 HPVISCACKTOLRVVNGIPRTFNIGMMWVSLRYRNKHIGGSLIKESWVLTAROCFPPSD 540
 QY 542 KDYEAMTGIHNVHGRGDEKCKQVLNVSQLVYGPESGDLVIMKLARPAVLDDFVSTIDLPN 601
 DB 542 KDYEAMTGIHNVHGRGDEKCKQVLNVSQLVYGPESGDLVIMKLARPAVLDDFVSTIDLPN 600

DB 541 KdYeaWlGhDvHnGrGdeKcKqVlnvsqLvygpegSdlvImklarpavlddfvstIdlpn 600
 QY 602 YGCTIPKTSQSVYGMWYTGILINDGLLRVAHLKYMGNKCSQHHRGKVTLENESEICAGA 661
 DB 601 YGCTIPKTSQSVYGMWYTGILINDGLLRVAHLKYMGNKCSQHHRGKVTLENESEICAGA 660
 QY 662 EKISGPGCEGDYGPPLVCEQHMKRMVLGVIYVPGGCAIPNRPGLFVRVAVYAKMIHKL 721
 DB 661 EKISGPGCEGDYGPPLVCEQHMKRMVLGVIYVPGGCAIPNRPGLFVRVAVYAKMIHKL 720
 QY 722 TYKVPQS 728
 DB 721 tykvyps 727

RESULT 14
 ID AAM88532 standard; protein: 728 AA.
 XX AAM88532:
 AC AAM88532:
 XX
 DT 02-MAR-1999 (first entry)
 XX
 DE Human hepatocyte growth factor (HGF) variant 4.
 XX
 KW Human; HGF: hepatocyte growth factor; variant; HGF receptor; medicine;
 XX heparan sulphate proteoglycan; cancer.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 91 /label= K91E
 FT /note= "wild-type Lys is replaced with Glu"
 FT Misc-difference 93 /label= K93E
 FT /note= "wild-type Lys is replaced with Glu"
 FT Misc-difference 94 /label= K94E
 FT /note= "wild-type Lys is replaced with Glu"
 XX
 PN W09851798-A1.
 XX
 PD 19-NOV-1998.
 XX
 PF 07-MAY-1998: 98WO-G801318.
 XX
 PR 10-MAY-1997: 97GB-0009453.
 XX
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
 XX
 PI Birchnmeier W, Cherrard E, Hartmann G;
 XX
 DR WPI: 1999-024213/02.
 XX
 XX Hepatocyte growth factor variants - binding to hepatocyte growth
 PT factor receptor, useful in medicine e.g. for cancer therapy
 PT
 PS Disclosure; Page -: 75pp; English.
 PS
 CC The present sequence represents a human hepatocyte growth factor (HGF)
 CC variant. The HGF variants of the invention bind to the HGF receptor, but
 CC are substantially unable to bind heparan sulphate proteoglycan. The HGF
 CC variants are useful in medicine e.g. variants having the same effects on
 CC target cells as wt-HGF can be administered to patients requiring HGF,
 CC while antagonistic variants can be administered to patients requiring HGF
 CC antagonists. They are especially useful for treating cancer. Some of the
 CC variants have a longer circulatory half-life in vivo and a greater
 CC mitogenic activity than wt-HGF in rats. They are therefore useful
 CC therapeutically as wt-HGF, but may have superior effects in vivo, e.g.
 CC give greater tissue penetration and ability to reach cell or tissue
 CC compartments.

CC Note: This sequence is not provided in the specification; it has been
CC created by modifying the wild-type HGF sequence.

SQ Sequence 728 AA;

Query Match	99.7%	Score 4113:	DB 20,	Length 728;
Best Local Similarity	99.6%	Pred. No. 7.5e-276;		
Matches 725; Conservative	2;	Mismatches 1;	Indels 0;	Gaps 0

QY	1	MMWKKLPA	LLHVVLLHLHLPLAT	YAEQQR	RNRNTI	HEFKSA	TKTLIK	TDPA	KIK	60
Db	1	mmwckllp	alllqvvllhlllpl	alyaeq	qrtrntlf	efksak	ttclik	tdpalk	ik	60
QY	61	TKKVN	ADOCAN	CTRNKGL	PFTCKA	EVFED	KARKQ	CLMP	FPNSMS	SGVKK
Db	61	tkkvnad	qcanctrn	kglpft	ckatfv	deae	eqclw	frnms	sgvkk	efn
QY	121	NKDYI	RNCI	ICKG	SGYK	GYVS	IT	SSG	IKCQ	PWSM
Db	121	nkdyirnc	iickgsgy	kvtsait	ksqik	cqgws	mldh	elsflps	rygkld	qen
QY	181	RGE	GGP	PCF	SN	EV	EV	CD	ICG	SE
Db	181	rgeegp	pcfsn	ev	ev	cd	icgse	ve	em	cln
QY	241	HRH	KFL	PER	Y	DK	EF	DD	N	Y
Db	241	hrhkfl	perydk	efddn	ycrnp	dgpr	pcy	cl	dp	ht
QY	301	ET	TE	IC	Q	G	G	E	G	Y
Db	301	eteciq	gqggy	rg	tv	cl	wn	gl	wp	cd
QY	361	ESP	PC	F	T	D	P	R	I	R
Db	361	espcft	dpri	rv	gcsq	lpc	dm	sh	gdcy	
QY	421	DL	AH	I	E	W	E	P	A	S
Db	421	dlahnie	wep	as	kl	n	en	y	c	r
QY	481	DH	P	I	S	C	A	K	T	Q
Db	481	dhyals	ca	k	t	q	l	r	v	ng
QY	541	L	K	O	E	A	M	I	G	I
Db	541	lkdyea	w	gl	nd	vn	gr	de	k	q
QY	601	NY	G	T	P	E	K	T	S	Y
Db	601	nygctli	pek	t	s	y	g	w	y	g
QY	661	AE	K	I	G	S	G	P	E	G
Db	661	aekit	gs	p	c	e	g	d	y	g
QY	721	L	T	K	V	P	O	S		
Db	721	ltykvp	ps	728						

Query Match	Best Local Similarity	99.7%	Score 4112;	DB 12;	Length 728;
Matches 726;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
QY 1	MWVYTLLEALLLQHLVLLHLLLPITAPPAEGQRKRRNTIHEKSKASTTLIKIPALAIK	60			
DB 1	mwvltlplalllqhlhlllllpiaipyaegyrkrntihetksaktlikidpalak	60			
OY 61	TKVYVTAOCANRCYRNKGLPFTCAFEYDARKOCLWPFENKSSGKKKEGHEFDLYE	120			
DB 61	tkkvntadqancrcrnnglpftcattrfdarkqcllfrfmsmgyskkkeghedfolye	120			
OY 121	NKDYTRNCITGKGRSYKGTVSTTKSGIKCOQWSSMIPEHSHFLPSSTYKGLQENYCRNP	180			
DB 121	nkdytrncitlgkgrsykgtvsttksgikcqpwsmliphehsflpsstrygkldgenycrnp	180			
OY 181	RGEEGPWCFTSNPEVAREVCIDIPCCSEVECTGCGESYRGIMDHTESGKICQRMDDHOTP	240			
DB 181	rgeegpwcftsnpevarevcidipccsevecmctcgesyrgimdtlesgkicqrdwhtp	240			
OY 241	HRHKFLPERYPDKGFDVNYCRNPDGQPPRPWCYTLTDPHTRWECYAIKTCADNTMNDTVPL	300			
DB 241	hrhkflperypdkgfdvnycrnpdgqpprpwcytltphterweycaiktcadntmndtvpl	300			
OY 301	ETTECTIOGEGYRGRTVMTINMGITPCQRMDSGX PHEHMTPEHFKCDLRENYCRNPDS	360			
DB 301	etctciogeggyrgtvtlmgipcqrdwsdgyphchmcpentfckdlrenycrnpdgs	360			
OY 361	ESPWCFETDPNIRVQYCSQIIPWCDSSHODCYRGNGKVMYMGTLSTGRGLTCSMMDXNME	420			
DB 361	espwcfetdprkivrgycsqipwcdsmshodcyrgngkvmynlsgtsglscsmdxnme	420			
OY 421	DLHRIHFWPPDASKINENYCRNPDDAHGPMVCTGNPLIPWDYCPISRCBGDTPTTYNL	480			
DB 421	dlnrlhfwppdaskinencrnpddahgpmvctgnplipwdycpisccegdtpcttynl	480			


```
OY 481 DHPVISC AKT KOLRVVNGIPTRTNIGMWSLARYRNKHICGSLIKESWLTAROCFPRS 540
    |||||||
Db 481 dhpvisca ktqlrvvnglptctnig mwslyrnkhicgslikeswltarqcipstrd 540
OY 541 LKDYEAMLGIHDYHGRGDEKCKQVLNVSQLVGPESGDLVLMKLRAPVLDDEVSTIDL 600
    |||||||
Db 541 lkdyeawlghdvhrrgdekckqvl nvsqlyvgsd lvmklarpavlddfvstidlp 600
OY 601 NYGCTIPEKTS CSVYGWGTGLINVDGLLRVAHLIYMGNEKCSQHHRGKVTLNSEI 660
    |||||||
Db 601 nygctipekts csvgwygylinydgl lrva hlyimgnekcsqhrgkvtlnseicag 660
OY 661 AEKIGSGPCGSDYGSPVCEQHQHMRWLVGIVPGRGCAIPNRPGEFVRVAVYAKWIKH 720
    |||||||
Db 661 aekigs gpcsgdy9sp l vceqhqhmrmvlgv lvpgrgcalpnrpg l fvravayakwihk 720
OY 721 LTYKVPQS 728
    |||||||
    ltykvpqs 728
```

Search completed: June 18, 2002, 17:46:59
Job time: 240 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2002, 17:45:19 ; Search time 24.39 Seconds

(without alignments)
2868.103 Million cell updates/sec

Title: US-09-423-516-2

Perfect score: 4126

Sequence: 1 MWVKRLPALDQHVHLHL.....VAYYAKWIKITLYKVPQS 728

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Processed: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4126	100.0	728	1 JH0579	hepatocyte growth
2	3804	92.2	728	1 A60185	hepatocyte growth
3	3802	92.1	728	1 A35644	hepatocyte growth
4	2806	68.0	710	1 I51283	hepatocyte growth
5	1821.5	44.1	411	1 I51285	hepatocyte growth
6	1652	40.0	711	1 A47136	macrophage-stimula
7	1615	39.1	716	1 JCS061	macrophage-stimula
8	1608	39.0	716	1 A40332	macrophage-stimula
9	1446.5	35.1	790	1 PLPG	plasma (EC 3.4.21
10	1408.5	34.1	812	1 PLBO	plasma (EC 3.4.21
11	1379.5	33.4	812	1 PLMS	plasma (EC 3.4.21
12	1371	33.2	810	1 PLHU	plasma (EC 3.4.21
13	1369	33.2	810	2 B30848	plasma (EC 3.4.21
14	1353.5	32.8	810	2 I46260	plasma (EC 3.4.21
15	1212.5	29.4	4548	1 S00657	apoptin (EC 3.4.21
16	1188.5	28.8	1420	2 A32869	apolipoprotein(a)
17	945	22.9	2869	2 T18518	apolipoprotein(a)
18	818	19.8	460	2 B61545	plasma (EC 3.4.21
19	784.5	19.0	455	2 A61545	plasma (EC 3.4.21
20	559.5	13.6	559	1 A35029	t-plasminogen acti
21	551.5	13.4	559	1 A29941	t-plasminogen acti
22	543	13.2	625	1 TBBO	thrombin (EC 3.4.2
23	532.5	12.9	562	1 UKHUT	t-plasminogen acti
24	531.5	12.9	618	2 A35827	thrombin (EC 3.4.2
25	527.5	12.8	622	1 TBHU	thrombin (EC 3.4.2
26	518	12.6	617	2 S10511	thrombin (EC 3.4.2
27	510.5	12.4	603	2 S28941	coagulation factor
28	480.5	11.6	655	1 A4688	hepatocyte growth
29	450	10.9	558	2 JCS878	plasma hyaluronan-

30	449	10.9	615	1 KFH012	coagulation factor
31	445	10.8	336	2 S33879	plasmin precursor
32	426.5	10.3	593	2 S45281	coagulation factor
33	423.5	10.3	560	1 JCA4795	plasma hyaluronan-
34	393	9.5	433	1 JN0560	u-plasminogen acti
35	390	9.5	431	1 UKHU	u-plasminogen acti
36	388	9.4	433	1 UKMS	u-plasminogen acti
37	386	9.4	442	1 UKPG	u-plasminogen acti
38	385	9.3	434	1 A35005	u-plasminogen acti
39	382.5	9.3	432	1 S18932	u-plasminogen acti
40	382.5	9.3	433	1 UKBAV	u-plasminogen acti
41	376.5	9.1	477	1 A43369	t-plasminogen acti
42	375	9.1	1035	1 A43090	enteropeptidase (E
43	373.5	9.1	477	2 JS0598	t-plasminogen acti
44	373	9.0	477	2 JS0597	t-plasminogen acti
45	372	9.0	273	2 A47246	trypsin (EC 3.4.2

ALIGNMENTS

RESULT 1
JH0579
hepatocyte growth factor precursor [validated] - human
N:Alternate names: hepatopoietin A; scatter factor
C:Species: Homo sapiens (man)
C>Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Dec-2000
C:Accession: JH0579; J00333; A41140; B36677; A36677; A33512; A39006; PH0114; A37796;
R:Seiki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
Gene 102, 213-219, 1991
A:Title: Organization of the human hepatocyte growth factor-encoding gene.
A:Reference number: JH0579; MUID:91340155
A:Accession: JH0579
A:Molecule type: DNA
A:Residues: 1-728 <SEK>
A:Cross-references: DDBJ:D90318
A:Note: The authors translated the codon GAA for residue 662 as Gly
R:Seiki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
Submitted to JIPID, March 1991
A:Description: Organization of the human hepatocyte growth factor-encoding gene.
A:Reference number: J00333
A:Accession: J00333
A:Molecule type: DNA
A:Residues: 1-481, 'RT', 484-728 <SE2>
R:Weidner, K.M.; Arakaki, N.; Hartman, G.; Vandekerckhove, J.; Weingart, S.; Rieder, Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991
A:Title: Evidence for the identity of human scatter factor and human hepatocyte growth factor.
A:Reference number: A41140; MUID:91334393
A:Accession: A41140
A:Molecule type: mRNA
A:Residues: 1-728 <ME1>
A:Cross-references: GB:M73239; NID:g337935; PIDN:AAA64239.1; PID:g337936
R:Seiki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya Biochem. Biophys. Res. Commun. 177, 321-327, 1990
A:Title: Isolation and expression of cDNA for different forms of hepatocyte growth fa
A:Reference number: A36677; MUID:91025062
A:Accession: B36677
A:Molecule type: mRNA
A:Residues: 1-728 <SE3>
A:Cross-references: GB:M60718; NID:g184031; PIDN:AAA52648.1; PID:g184032
A:Accession: A36677
A:Molecule type: mRNA
A:Residues: 1-161, 167-728 <SE4>
A:Cross-references: EMBL:X16322
A:Experimental source: Leukocyte
R: Miyazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nak Biochem. Biophys. Res. Commun. 163, 967-973, 1989
A:Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth
A:Reference number: A33512; MUID:89392017
A:Accession: A33512
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-728 <MY>

A:Cross-references: GB:M29145; NID:g184041; PIDN:AA52650.1; PID:g306846
R:Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hilt
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
A:Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocyte
A:Reference number: A39006; MUID:9110540
A:Accession: A39006
A:Molecule type: mRNA
A:Residues: 1-161,167-728 <RUB>
A:Cross-references: GB:M55379
A:Experimental source: embryonic lung
R:Toshiyama, Y.; Aizaki, N.; Naka, D.; Takahashi, K.; Hirose, S.; Kondo, J.; Nakayama,
Biochem. Biophys. Res. Commun. 175, 660-667, 1991
A:Title: Identification of the N-terminal residue of the heavy chain of both native and
A:Reference number: PH0114; MUID:91207365
A:Accession: PH0114
A:Molecule type: protein
A:Residues: 32-43;53-58 <YOS>
A:Experimental source: plasma
R:Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.
J. Cell Biol. 111, 2097-2108, 1990
A:Title: Scatter factor: molecular characteristics and effect on the invasiveness of epi
A:Reference number: A37796; MUID:91035621
A:Accession: A37796
A:Molecule type: protein
A:Residues: 86-91;329-344;356-363, 'XX', 366-370;425-434;442-447, 'X', 449-450;543-546, 'X', 5
R:Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimonishi, M.; Sugimura, A.; Tash
Nature 342, 440-443, 1989
A:Title: Molecular cloning and expression of human hepatocyte growth factor.
A:Reference number: S06794; MUID:90066676
A:Accession: S06794
A:Molecule type: mRNA
A:Residues: 1-31, 'HK', 34-77, 'N', 79-292, 'V', 294-299, 'M', 301-316, 'A', 318-335, 'K', 337-386, '
A:Cross-references: EMBL:X16323; NID:g32081; PIDN:CAA34387.1; PID:g32082
A:Experimental source: liver
A:Note: the authors translated the codon CAG for residue 727 as Glu
A:Note: part of this sequence, including the amino end of both the alpha and beta chains
R:Hartman, G.; Naldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchm
Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth fact
A:Reference number: I59214; MUID:93087571
A:Accession: I59214
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-288, 'ET' <HAR>
A:Cross-references: GB:I02931; NID:g184033; PIDN:AA52649.1; PID:g184034
R:Myazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
Eur. J. Biochem. 197, 15-22, 1991
A:Title: An alternatively processed mRNA generated from human hepatocyte growth factor g
A:Reference number: S15443; MUID:91200041
A:Accession: S15443
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-288, 'ET' <MY2>
A:Cross-references: EMBL:X57574; NID:g32083; PIDN:CAA40802.1; PID:g32084
R:Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.
Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991
A:Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: cloning
A:Reference number: I52253; MUID:92062058
A:Accession: I52253
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 161-166 <SHI>
A:Cross-references: GB:S62561; NID:g237996; PIDN:AA520169.1; PID:g237997
A:Accession: S62561
A:Gene: GDB:HGFP
A:Cross-references: GDB:127524; OMIM:142409
A:Map position: 7921.1-7921.1
A:Intons: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 469/1; 48
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C:Function:
A:Description: stimulates mitosis of hepatocytes and other cells
A:Note: does not have proteinase activity
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology

C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle;
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-494,495-728/Product: hepatocyte growth factor #status experimental <MAT>
F:32-494/Domain: alpha chain #status experimental <ACH>
F:128-206/Domain: kringle homology <KR1>
F:211-288/Domain: kringle homology <KR2>
F:305-383/Domain: kringle homology <KR3>
F:391-469/Domain: kringle homology <KR4>
F:495-728/Domain: beta chain #status experimental <BCH>
F:495-716/Domain: trypsin homology <TRY>
F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status expert
F:294,402,566,653/Binding site: carboxylate (Asn) (covalent) #status predicted
F:487-604/Disulfide bonds: #status predicted

Query Match 100.0%; Score 4126; DB 1; Length 728;
Best Local Similarity 100.0%; Pred. No. 2,4e-282;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MMVTKLPLALLQHVLLHLPLAIPYAEQRRRRNTHEFKSAKTLIKIDPALKIK 60
1 MMVTKLPLALLQHVLLHLPLAIPYAEQRRRRNTHEFKSAKTLIKIDPALKIK 60
1 TKKVTADQCANRGTNRKGLPFTCKAFYFDKARQCLMFPPNSSSGYKKKEGHEFDLYE 120
1 TKKVTADQCANRGTNRKGLPFTCKAFYFDKARQCLMFPPNSSSGYKKKEGHEFDLYE 120
1 NKDYIRNCTIIGKRSYKSTVITKSGIKCOPWSSMIPHEHFEFLSSYKGLQENYCNRP 180
1 NKDYIRNCTIIGKRSYKSTVITKSGIKCOPWSSMIPHEHFEFLSSYKGLQENYCNRP 180
1 RGEEGPWCFTSNPEVRYEVDIPQCESEVEMTCNGESYRGLMDHTEGKICQRMHDQTP 240
1 RGEEGPWCFTSNPEVRYEVDIPQCESEVEMTCNGESYRGLMDHTEGKICQRMHDQTP 240
1 HRHFLPRYTDKGFDDNYCNRPDGPDPKCYTLDPHRMVECAIKTCADTMDTVPPL 300
1 HRHFLPRYTDKGFDDNYCNRPDGPDPKCYTLDPHRMVECAIKTCADTMDTVPPL 300
1 ETEECIOGQGGEGYGTNTINNGIPPCQRMDSQYRPHENMPENFEKCDLRENYCRNPDGS 360
1 ETEECIOGQGGEGYGTNTINNGIPPCQRMDSQYRPHENMPENFEKCDLRENYCRNPDGS 360
1 ETEECIOGQGGEGYGTNTINNGIPPCQRMDSQYRPHENMPENFEKCDLRENYCRNPDGS 360
1 ESPMCFETDPIRNVGYCSQIPNCDMSHGQDCYRGNGKYNMGNLSQTRSGLTCSMMDKME 420
1 ESPMCFETDPIRNVGYCSQIPNCDMSHGQDCYRGNGKYNMGNLSQTRSGLTCSMMDKME 420
1 DLHRIIFWEPDASKLNEYCNRPDDDAHGPCYTGNPILPMDYCPISRCBEDTPTTYNL 480
1 DLHRIIFWEPDASKLNEYCNRPDDDAHGPCYTGNPILPMDYCPISRCBEDTPTTYNL 480
1 DHPVTSCKATKQLRVNGIPRTNIGWVSLRYRNKHKICGSLIKESVNLTAAROCFSPSRD 540
1 DHPVTSCKATKQLRVNGIPRTNIGWVSLRYRNKHKICGSLIKESVNLTAAROCFSPSRD 540
1 LKDYEAWLGIHDVHGRGDEKCKOVLNVSQLYVGPESDLVLMKLAAPVLDPEVSTDLPL 600
1 LKDYEAWLGIHDVHGRGDEKCKOVLNVSQLYVGPESDLVLMKLAAPVLDPEVSTDLPL 600
1 NYGCTIPKRTSCSYVGMWYTGILINDGLLRVAHLTYMNEKCSQHHRRKVTLNSEICAG 660
1 NYGCTIPKRTSCSYVGMWYTGILINDGLLRVAHLTYMNEKCSQHHRRKVTLNSEICAG 660
1 AEKISGCEGEGYGLVCEQHKMVMVGLVYVPGSCAIPNRPGLFVAVATYAKWIIHII 720
1 AEKISGCEGEGYGLVCEQHKMVMVGLVYVPGSCAIPNRPGLFVAVATYAKWIIHII 720
1 LTYKVPQS 728
1 LTYKVPQS 728

RESULT 2

A60185
 hepatocyte growth factor precursor - mouse
 N:Alternate names: hepalectin A: scatter factor
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Mar-1993 #sequence-revision 26-May-1994 #text-change 16-Jun-2000
 C:Accession: J02117; PC2064; A60185; S43416; S45521; S17173; S10966; I48758; J00231
 R:Sasaki, M.; Nishio, M.; Sasaki, T.; Enami, J.
 Biochem. Biophys. Res. Commun. 199, 772-779, 1994
 A:Title: Identification of mouse mammary fibroblast-derived mammary growth factor as hepalectin
 A:Reference number: J02117; M0ID:94183257
 A:Accession: J02117
 A:Molecule type: mRNA
 A:Residues: 1-728 <SA2>
 A:Cross-references: GB:D10212; NID:q220435; PIDN:BA01064.1; PID:q220436
 A:Experimental source: fibroblast, COS-1 cell
 A:Note: submitted to JIPID, May 1993
 A:Accession: PC2064
 A:Molecule type: protein
 A:Residues: 496-504 <SA2>
 A:Cross-references: E.M.: Meromsky, L.; Setier, E.; Winter, D.W.; Goldberg, I.D.
 Proc. Soc. Exp. Biol. Med. 195, 34-43, 1990
 A:Title: Purified scatter factor stimulates epithelial and vascular endothelial cell migration
 A:Reference number: A60185; M0ID:90377927
 A:Accession: A60185
 A:Molecule type: protein
 A:Residues: 'X', 184-188, 'KX', 191-192, 'X', 194, 'XX', 197, 357-364, 'XX', 367, 375-377, 'E', 379, 'R', 380
 R:Li, Y.; Michalopoulos, G. K.; Zarnegar, R.
 Biochim. Biophys. Acta 1216, 299-303, 1993
 A:Title: Molecular cloning and characterization of cDNA encoding mouse hepatocyte growth factor
 A:Reference number: S43416; M0ID:94060105
 A:Accession: S43416
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-728 <LID>
 A:Cross-references: EMBL:X72307
 R:Li, Y.
 submitted to the EMBL Data Library, May 1993
 A:Reference number: S45521
 A:Accession: S45521
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-563, 'H', 565-728 <L12>
 A:Cross-references: EMBL:X72307
 R:Coffier, A.; Fellows, J.; Young, S.; Pappin, D.; Rahman, D.
 Biochem. J. 278, 35-41, 1991
 A:Title: Purification and characterization of biologically active scatter factor from rat liver
 A:Reference number: S17173; M0ID:91354223
 A:Accession: S17173
 A:Molecule type: protein
 A:Residues: 496-517, 'T', 519 <COF>
 R:Gherardi, E.; Stoker, M.
 Nature 346, 228, 1990
 A:Title: Hepatocytes and scatter factor.
 A:Reference number: S10966; M0ID:90326152
 A:Accession: S10966
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 496-507, 'X', 509-512, 'L', 514-516, 'X', 518-519 <NAT>
 R:Plaschke-Schluter, A.; Behrens, J.; Gherardi, E.; Birchmeier, W.
 J. Biol. Chem. 270, 830-836, 1995
 A:Title: Characterization of the scatter factor/hepatocyte growth factor gene promoter.
 A:Reference number: I48758; M0ID:95122532
 A:Accession: I48758
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-30 <RES>
 A:Cross-references: EMBL:X81630; NID:9673451; PIDN:CA57286.1; PID:9673452
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Function: stimulates mitosis of hepatocytes and other cells
 A:Note: does not have proteinase activity
 C:Superfamily: hepatocyte growth factor; kringe homology; trypsin homology
 C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringe; pyz

F:1-32/Domain: signal sequence #status predicted <SIG>
 F:56-495/Domain: hepatocyte growth factor #status predicted <MAT>
 F:56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
 F:129-207/Domain: kringe homology <KR1>
 F:212-289/Domain: kringe homology <KR2>
 F:306-384/Domain: kringe homology <KR3>
 F:392-470/Domain: kringe homology <KR4>
 F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
 F:496-719/Domain: trypsin homology <TRY>
 F:733/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
 F:295,403,569,656/Binding site: carbonyl (asn) (covalent) #status predicted
 F:488-607/Disulfide bonds: #status predicted

Query Match	92.2%	Score 3804	DB 1	Length 728
Best Local Similarity	90.8%	Pred. No. 1e-259		
Matches 660	Conservative 39	Mismatches 26	Indels 2	Gaps 1
1	MMVTKLPLALLQHVLLHLLLPALPAIPAEQRRRRNTHEFKSAKTLIKIDPALKIK	60		
2	MMGTKLPLVLLQHVLLHLLHVAIPAEQRRRRNTHEFKSAKTLIKIDPALKIK	61		
61	TKKVNADQCAKPTTRNKGDPFTCKAFVFDKARKQCLMPPNSMSGVKKFGEHFDLYE	120		
62	TKKNSADECANPTTRNKGDPFTCKAFVFDKARKQCLMPPNSMSGVKKFGEHFDLYE	121		
121	NKQVIRNCTIKGRSYKQVTSITSGIKCQWSSMIPHEHFLSSYKGLQENYCNRP	180		
122	NKQVIRNCTIKGRSYKQVTSITSGIKCQWSSMIPHEHFLSSYKGLQENYCNRP	181		
181	RGEFGPWCFTSNPEVREVEDIPQCEVEECMTNGESRGLMDHTESGICQWMDHOTP	240		
182	RGEFGPWCFTSNPEVREVEDIPQCEVEECMTNGESRGLMDHTESGICQWMDHOTP	241		
241	HRHKLPERYDPKGFDNDYCNPDGQPRPWCYTLDPTRMWCAIKTCADNTMNDVPL	300		
242	HRHKLPERYDPKGFDNDYCNPDGQPRPWCYTLDPTRMWCAIKTCADNTMNDVPL	301		
301	ETTECIGQSGYRGYVNTINGIIPQWRMDQYHEHDMTEENKCKDLRENYCNRPDS	360		
302	ETTECIGQSGYRGYVNTINGIIPQWRMDQYHEHDMTEENKCKDLRENYCNRPDS	361		
361	ESPMCFETDPNIRVGYCQIIPNCDSHGQDCYRNGKRYMGNLSOTRSLGCSMMKME	420		
362	ESPMCFETDPNIRVGYCQIIPNCDSHGQDCYRNGKRYMGNLSOTRSLGCSMMKME	421		
421	DLHRHIFWEPDASKLNNYCNPDGQPRPWCYTLDPTRMWCAIKTCADNTMNDVPL	480		
422	DLHRHIFWEPDASKLNNYCNPDGQPRPWCYTLDPTRMWCAIKTCADNTMNDVPL	481		
481	DHPVISCATKQLRVNGIIPRTNIGMNVSLRYRNKHTCGSSLKESNVLTAROCFPR	539		
482	DHPVISCATKQLRVNGIIPRTNIGMNVSLRYRNKHTCGSSLKESNVLTAROCFPR	541		
540	DLKDYEALGILHNVHGGDEKCOVLNVOLVYGPESGSDLVLMKLARPAVLDFVSTID	598		
542	KDLKDYEALGILHNVHGGDEKCOVLNVOLVYGPESGSDLVLMKLARPAVLDFVSTID	601		
599	LPNVGTIPEKTSYGVGWTGLINVDGLLRVAHLTYMGNEKCSQHHRGKVTINSEIC	658		
602	LPNVGTIPEKTSYGVGWTGLINVDGLLRVAHLTYMGNEKCSQHHRGKVTINSEIC	661		
659	AGAERIGSGPESGSDLVLMKLARPAVLDFVSTID	718		
662	AGAERIGSGPESGSDLVLMKLARPAVLDFVSTID	721		
719	ITLTVKV 725			
722	VILTVKL 728			

RESULT 3
 A35644

```

Db      362  ESPMCTTTPNINVGYSQIPIKCDVSSGDCYRGKNKMYGNLSKTRSLGTCSMMDKNE 421
Oy      421  DLRRHIFWEPDASKLBNENCRNPDDAHGPMCTGSPDLLPMYCPISRCRGOTPIIVNL 480
Db      422  DLRRHIFWEPDASKLKRNICRNPDDAHGPMCTGSPDLLPMYCPISRCRGOTPIIVNL 481
Oy      481  DHEVISCATKQOLRVVNGIPTRTNIGMWSLRBYRNHIGSGSLIKESWLTAROGFSPSR 539
Db      482  DHEVISCATKQOLRVVNGIPTRTVGMWVSLKRYRNHIGSGSLIKESWLTAROGFSPRAN 541
Oy      540  -DLKDYEAUWGLIHVDYGRDEKCKQYLNLSQLYGREGSDVLMKLARAVLDDPYSTID 598
Db      542  KDLKDYEAUWGLIHVDYGRDEKCKQYLNLSQLYGREGSDVLMKLARAVLDDPYSTID 601
Oy      599  LPNVGCTIEKTSCTSYVGMGYTGLINVDGLLRVANHLYINGNEKSCSOHHGKTYLNESEC 658
Db      602  LPNVGCTIEKTSCTSYVGMGYTGLINADGLLRVANHLYINGNEKSCSOHHGKTYLNESEC 661
Oy      659  AGAEKIGSGPCEDYGGPLVCEQHKMRYLVGIYVDEGSCAIDNRPQIFRVAYYAKWIMK 718
Db      662  AGAEKIGSGPCEDYGGPLVCEQHKMRYLVGIYVDEGSCAIDNRPQIFRVAYYAKWIMK 721
Oy      719  ILITKYV 725
Db      722  VILITKYL 728

RESULT 4
151283
hepatocyte growth factor precursor - clawed frog
N.Alternate names: hepatolefin A; scatter factor
C.Species: Xenopus sp. (clawed frog)
C.Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C.Accession: I51283
R.Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiohawa, K.
Mech. Dev. 49, 123-131, 1995
A.Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in Xenopus
A.Reference number: I51283; MUID:95267690
A.Accession: I51283
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-710 <NAA>
A.Cross-references: GB:S77422; NID:g998933; PIDN:ABAB34354.1; PID:g998933
A.Note: the authors' translation for residue 458 (Thr) is inconsistent with the nucle
C.Function: disulfide-bonded heterodimer of chains derived from the same precursor
A.Description: stimulates mitosis of hepatocytes and other cells
A.Note: does not have proteinase activity
C.Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C.Keywords: duplication; glycoprotein; growth factor; heterodimer; kringle
F:42-477/478-709/Product: hepatocyte growth factor #status predicted <MAT>
F:42-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F:115-193/Domain: kringle homology <KR1>
F:198-275/Domain: kringle homology <KR2>
F:289-367/Domain: kringle homology <KR3>
F:375-453/Domain: kringle homology <KR4>
F:478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
F:478-700/Domain: trypsin homology <TRY>
F:52-128/281,322,379,550,637,666/Binding site: carbohydrate (Asn) (covalent) #status
F:470-588/Disulfide bonds: #status predicted

Query Match 68.0%; Score 2806; DB 1; Length 710;
Best Local Similarity 66.1%; Pred. No. 1,5e-189;
Matches 481; Conservative 107; Mismatches 120; Indels 20; Gaps 5;

Oy      1  MMVVKLLPALLQNVHLHLPLTAIPAEQORRRNTIHEFKSATTLLIKIDPAIKIK 60
Db      1  MMVTKQWVIFL-----ITLAEGRGKRRNAFDYKKTAAETTLRLKALEVK 47
Oy      61  TRKVTADQCANRGTNRNGLPFTCKAFEDFARQCLMFPNSMSGVKKKEFGHEPFLYE 120

```

```

Db 48 TKMNTTENCARKSRNKGLPETCKAFAPDKNIKRCHEWSEFNTMSAGIKDKDYIDSPDLYE 107
OY 121 NKDYINRCITIGKRSYKGVSTITKSGIKCOPMSSMTPHESFLPSSYRGKDLOENCRNP 180
Db 108 KQDYINDCJHGGSNRGRNTKRLAQCPNMSMPHEHSFLPSYRGDLKENCRNP 167
OY 181 RGESEGPWCFTSPREYEVCDIPQSEVEMCMNGESYGLMDHTESGKICQRMHDQRP 240
Db 168 KGESEGPWCFTSPREYEVCDIPQSEVEMCMNGESYGLMDHTESGKICQRMHDQRP 227
OY 241 HRHKLPERYPDKGFDDNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTMNDTVPL 300
Db 228 HKHKEFRPERYPDKGFDDNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTMNDTVPL 285
OY 301 ETEECIOGEGYRGVNTWNGICQRMDSQYRPHEDMTPEHFCKKDLRENTCRNPDGS 360
Db 286 TKDCCKGGEGYRGVNTWNGICQRMDSQYRPHEDMTPEHFCKKDLRENTCRNPDGS 344
OY 361 ESPWCFTTDPNIRIVGCSQIPNCDMSHGDCYRGNGKNYNGNLSQTRSGLTCSMDPKNM 420
Db 345 ESPWCFTTDPNIRIVGCSQIPNCDMSHGDCYRGNGKNYNGNLSQTRSGLTCSMDPKNM 404
OY 421 DHRHITFEPDASKLNTNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTMNDTVPL 480
Db 405 DHRHITFEPDASKLNTNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTMNDTVPL 464
OY 481 DHRHITFEPDASKLNTNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTMNDTVPL 539
Db 465 DHRHITFEPDASKLNTNCRNPDGQRPWCYTLDPHTWEXCAIKTCADNTMNDTVPL 523
OY 540 DLKYEAMLGSHVGRDECKQVLANYSOLVYEGSDLYLMDKARPAVLDDPFTSID 598
Db 524 DLKYEAMLGSHVGRDECKQVLANYSOLVYEGSDLYLMDKARPAVLDDPFTSID 562
OY 599 LRYNGCTTPEKISCSYGGYGLNYDGLRYAHLYIMGNEKSCQHRGKVTLNESEIC 658
Db 583 LRYNGCTTPEKISCSYGGYGLNYDGLRYAHLYIMGNEKSCQHRGKVTLNESEIC 642
OY 659 ACAERKISGPGCGDYGPGVPCGSHKMWLYGVIPGRGALNPRGPIFRVAYAAWIHK 718
Db 643 ALGFTANIPCEPBERDYGGLICEBNKTHLVQVYIIPGRGALNPRGPIFRVAYAAWIHK 702
OY 719 ILTYKVP 726
Db 703 IMLTYKAP 710

RESULT 5
151285
hepatocyte growth factor/scatter factor - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: 151285
R:Streit, A.; Stern, C.D.; Thery, C.; Ireland, G.W.; Aparicio, S.; Sharpe, M.J.; Cheraud
Development 121, 813-824, 1995
A:Title: A role for HGF/SF in neural induction and its expression in Hensen's node durin
A:Reference number: 151285; MUID:9537013
A:Accession: 151285
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-411 <STR>
A:Cross-references: GB:S77480; NID:9998675; PID:9998676
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
F:124-197/Domain: kringle homology <KR3>
F:202-279/Domain: kringle homology <KR2>
F:296-374/Domain: kringle homology <KR3>

Query Match 44.1%; Score 1821.5; DB 2: Length 411;
Best Local Similarity 73.1%; Pred. No. 14e-120;
Matches 309; Conservative 49; Mismatches 52; Indels 9; Gaps 2;

OY 1 MMVTKLLPALLQHVLLHLLPIAIPVNEGQRKRNTIHEFKSAKTLIKIDPALKIK 60

```



```

158 YCRNPDNDENPCMYTDDPDKRYDYCDIPED-EDKCMHCSEGENEGKIAKTMGSRDCQAM 216
OY 236 DHQPHRHKFLPERYPKGFNDNCRNPDGPRRWCYTLDDHTREYCAITCADNTMND 295
Db 217 DQSPHAGYIPSPKPNKLMNCRNPDGPRRWCYTLDDHTREYCAITCADNTMND 270
OY 296 TDVPLE--TTTCIGOGEGYRGVNTIWNIGIPCRMDQYHEHDMPEENKCKDLRENY 353
Db 271 TPRPSSGKCYCLGKGTGKNYGYAVYTESGHTCQRMSEQTTHKHNKTRTENPCKNLENY 330
OY 354 CRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSH-----GQDCYRNG 396
Db 331 CRNPDGSESPWCFTTDPNIRVGYCSQIPNCDMSH-----GQDCYRNG 389
OY 397 KNYGNLSQTSGLTCSMW-----DXN----- 418
Db 390 QSYRGTSTYITIGRKCOSMSMTPHRLKTPENPNAGLTNNYCRNPDADKSPWCYTDP 449
OY 419 ----- 418
Db 450 RVRRPFCMLKACSETPPEQVPAPOAPGVENPPEADCMIGTKSKRGKATTVAGVPCQEW 509
OY 419 -MEDLHRIHEWEPDA---SKLNENYCRNPDGPRRWCYTLDDHTREYCAITCADNTMND 474
Db 510 AAGRPQHOSITPFTNPDQSGLENYCRNPDGPRRWCYTLDDHTREYCAITCADNTMND 567
OY 475 PTYNLDPHYVSCAKTK-----OLRVNGIPIPTNT-GMWSLRYRNKHGGLSGLIES 527
Db 568 -----FDCKRPKEPKKSGRIYGGVCSKPSHWPQVSLERSRHGCGGLLSPK 617
OY 528 WVLTAAROCFSP-RDLKDEYAMLGIHDVHGRGDEKCOVLANYSOLVYSGESGLVLMKLAR 586
Db 618 WVLTAAROCFSP-RDLKDEYAMLGIHDVHGRGDEKCOVLANYSOLVYSGESGLVLMKLAR 675
OY 587 PAVLDDEFTYIDLPNYGTCTPEKTSVYGVGYGLIYNDGLRLVLAHYLYINGNEKCSOH 646
Db 676 PAITKEVYIPACLDPRPNVMAARTECYITGMGEGTGTGEGILKAHLVTEKNKCNNE 735
OY 647 RGKTLNESELCAAGAKETGSGPCEDYDGLVCEQHKRMVLCVYPRGCAIPRPGIE 706
Db 736 YLDRVAKRTELCAGHLIGTDSGCGDSGGLVCEPEKRYIIQGVYISWGLPRPKPGVY 795
OY 707 VRVAYYAKMIHK 718
Db 796 VRVAYYAKMIHK 807

RESULT 11
PLMS
Plasmin (EC 3.4.21.7) precursor - mouse
N:Contains: angiotatin; plasminogen
C:Species: Mus musculus (house mouse)
C:Date: 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
C:Accession: A38514; S48202; S48203
R:Degeen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
Genomics 8, 49-61, 1990
A:Title: Characterization of the cDNA coding for mouse plasminogen and localization of
A:Reference number: A38514; MUID:91184812
A:Accession: A38514
A:Molecule type: mRNA
A:Residues: 1-812 <DEG>
A:Cross-references: GB:J04766; NID:g200402; PID:AAA50168.1; PID:g200403
R:Lijnen, H.R.; Van Hoet, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MUID:95010076
A:Accession: S48202
A:Molecule type: protein
A:Residues: 20-25 <LIU>
A:Accession: S48203
A:Molecule type: protein
A:Residues: 22-27 <LI2>
```

```

C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many
C:Comment: Plasminogen is converted into plasmin by plasminogen activators, both plas
C:Comment: immediately after dissociation from the clot. In the presence of the inhibitor, the act
C:Comment: inhibitor, the activation involves also removal of the activation peptide.
C:Comment: Strimelysin I (see PIR:KCMSI) acts on plasminogen to produce angiotatin.
C:Function:
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C:Keywords: angiotensin inhibitor; blood; duplication; fibrinolysis; glycoprotein; h
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-812/Product: plasminogen #status predicted <PRO>
F:20-96/Domain: activation peptide #status predicted <APT>
F:79-466/Product: angiotatin #status predicted <AST>
F:97-581,582-812/Product: plasmin #status predicted <MAT>
F:103-181/Domain: chain A #status predicted <ACH>
F:185-262/Domain: kringle homology <KR1>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:582-812/Domain: chain B #status predicted <BCB>
F:582-805/Domain: trypsin homology <TRY>
F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335
bonds: #status predicted
F:78-79/Cleavage site: Glu-Asn (stromelysin I) #status predicted
F:136,308/Binding site: carboxylate (Asn) (covalent) #status predicted
F:466-467/Cleavage site: Thr-Val (plasminogen I) #status predicted
F:581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
F:624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 33.4%; Score 1379.5; DB 1; Length 812;
Best Local Similarity 34.4%; Pred. No. 3.6e-89;
Matches 288; Conservative 119; Mismatches 262; Indels 167; Gaps 22;

OY 15 VLLHLLPLIAIPVDEGOKRRNTIHERKSAKTLIKIDPLAKIKTKVNTA---DOCA 71
Db 7 ILLFLL-----RPGQ---GDSLDGYISTOGASLFSK-----TKQLAAGVSDCL 50.
OY 72 NRCTRNGKLPFTCKAFVDFPKARKOCLMPFPNSMGVKKPFGEHFDYENKRYINCIIG 131
Db 51 AKCEBE--TDFVCRSFQYHSKQOCVIAENSKTSITM--RDVYLEKRYLSECKTG 106
OY 132 KGRSYKGVYITKSGIKCPWSSMIPHEHSFLPSSYRGKLOENYCRNPDGEGPWCFT 191
Db 107 IONGYRGIMSKRKSQVACQKWAATPRHVNYSPRHNGLEENYCRNPDNDGQPMCT 166
OY 192 SNPEVREYECIDIPGSEVCEMTONGESYRGMDHTESGKTCORMDHQPHRHKFLPERYP 251
Db 167 TDPDRKRYDCNIPCEBE--ECMYCSEKEYEGKISKTSMDGDCAMPQSOPSHAGYIPAKFP 225
OY 255 DKGFDNDCRNPDGPRRWCYTLDDHTREYCAITCADNTMNDVPLETTEGIGOGGE 311
Db 226 SNLKNKNCCHNDGPRRWCYTLDDHTREYCAITCADNTMNDVPLETTEGIGOGGE 281
OY 312 GYRGTVNTIWNIGIPCRMDQYHEHDMPEENKCKDLRENYCRNPDGSESPWCFTTDPN 371
Db 282 NRGTVSVTVSGKTCQRMSEQTTHKHNKTRTENPCKNLEENYCRNPDGTAATTCITDSQ 341
OY 372 IRVGYCSQIPNCDMSHG-----ODCYRNGKNYGNLSQTSGLTCSM 414
Db 342 LRMEYC-ETPSCSSASPQSDSVPRPEQTPVYECYOSDQSGYSGTITGKKCS 400
OY 415 WDKNMDLHRIHFW---PDASKLNENYCRNPDGPRRWCYTLDDHTREYCAITCADNTMND 471
Db 401 WAMAFR--HRHSKTPENFPDAG-LENNYCRNPDG--KGMWCYTTPSVMEYCNLKR-S 455
OY 472 DTPPTIVNLDPHYVSCAKTK----- 486
```

Db 456 ETGSGVEL--PTVSEPSPSDSEDCMYGNGKDYRGKTAVTAAQTGCGMAQEPHRH 513
 QY 487 -----CAKTKL----- 493
 Db 514 SIPTPTNPRADELEKNYCRNPDPGDVNGPWCYTTNPRKLYDCDIPLCASASSFECKRPQV 573
 QY 494 -----RVVNG-IPRTNIGMVMVSLRYR--NKHICGSLIKESWVLTFRQFSPRDLKD 543
 Db 574 EPKCPGKAVYGGCVANPMSWPQISLRTRFTGQHCGLLPAENVLTARHLEKSSRRE 633
 QY 544 -YEAMVLGHDVHGRCDEKCKOVLNVSQLVYGPESGDLVLMKLARPAVLDEPVSTIDLPNY 602
 Db 634 FYKVLGAHEEYIRGLD--VOEISVAKLILEPNNRDIALKLSPATITDKVIAPACLPSP 691
 QY 603 GCTIPEKTSQSYGMYGGLINVDGLLVAHLIYMGNEKSCSHHGKATLNSSEICAGNE 662
 Db 692 NYMADRTICYTGGTGTGFTGACRLAEOLPVLENKVCNVEYLNNRKSTELCAQOL 751
 QY 663 KIGSGPCGDDYGGPLVCEQHMKRMVLGYIVPGGCAIPNPGIFVYVAVYAKMKH 718
 Db 752 AGVDSCGSGDGGPLVCEFKDKYILQGVTSNGLGCAIRPKRPSVYRVSRFVMIER 807
 RESULT 12
 PLHU
 Plasmin (EC 3.4.21.7) precursor [validated] - human
 N:Alternate names: plasminogen precursor [misnomer]
 N:Contents: angiotatin, micropasmin, plasminogen
 C:Species: Homo sapiens (man)
 C:Date: 24-Apr-1984 #sequence_revision 02-Dec-1994 #text_change 15-Sep-2000
 C:Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A04625; A04
 J:Peterzen, T.E.; Martzen, M.R.; Ichinose, A.; Davle, E.W.
 J: Biol. Chem. 265, 6104-6111, 1990
 A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr
 A:Reference number: A35229; MUID:90202879
 A:Accession: A35229
 A:Molecule type: DNA
 A:Residues: 1-810 <PEP>
 A:Cross-references: GB:J05286; GB:M34276; NID:q190064; PIDN:AAA60113.1; PID:g387026
 R:Margaret, N.; Bruno, L.; Pontoglio, M.; Candian, G.; Meroni, G.; Ottolenghi, S.; Ta
 Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990
 A:Title: Definition of the transcription initiation site of human plasminogen gene in 11
 A:Reference number: I52242; MUID:91097523
 A:Accession: I52242
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-16 <MAL1>
 A:Cross-references: GB:M62890; NID:q190092; PIDN:AAA36454.1; PID:g553613
 R:Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.
 J. Biol. Chem. 263, 254-260, 1987
 A:Title: Molecular cloning and characterization of a full-length cDNA clone for human pl
 A:Reference number: A26646; MUID:87162490
 A:Accession: A26646
 A:Molecule type: mRNA
 A:Residues: 1-471, 'D', 473-810 <FOR>
 A:Cross-references: GB:J05199; NID:g35530; PIDN:CAA28831.1; PID:g35531
 R:Experimental source: liver
 R:Malinowski, D.P.; Sadler, J.E.; Davle, E.W.
 Biochemistry 23, 4243-4250, 1984
 A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and
 A:Reference number: I45961; MUID:8502311
 A:Accession: I62738
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 292-471, 'D', 473-810 <MAL2>
 A:Cross-references: GB:K02922; NID:q190112; PIDN:AAA60124.1; PID:g387031
 A:Accession: I84609
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 367-419 <MAL3>
 A:Cross-references: GB:K02921; NID:q190110; PIDN:AAA60123.1; PID:q190111
 R:Brumsholtz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.

Eur. J. Biochem. 114, 465-470, 1981
 A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of huma
 A:Reference number: S03735; MUID:81212097
 A:Accession: S03735
 A:Molecule type: protein
 A:Residues: 20-71, 'E', 73-76 <BRU>
 R:Sortrup-Jensen, L.; Petersen, T.E.; Magnusson, S.
 submitted to the Atlas, July 1977
 A:Reference number: A00929
 A:Accession: A00929
 A:Molecule type: protein
 A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>
 R:Wiman, B.
 Eur. J. Biochem. 76, 129-137, 1977
 A:Title: Primary structure of the B-chain of human plasmin.
 A:Reference number: A04627; MUID:77225245
 A:Accession: A04627
 A:Molecule type: protein
 A:Residues: 581-810 <WIL>
 R:Wiman, B.; Wallen, P.
 Eur. J. Biochem. 50, 489-494, 1975
 A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human
 A:Reference number: A04625; MUID:75093329
 A:Accession: A04625
 A:Molecule type: protein
 A:Residues: 20-50, 'Q', 51-71, 'E', 73-85, 87-100 <WIZ>
 R:Wiman, B.; Wallen, P.
 Eur. J. Biochem. 58, 539-547, 1975
 A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen
 A:Reference number: A04626; MUID:76043692
 A:Accession: A04626
 A:Molecule type: protein
 A:Residues: 483-507, 'E', 509-604 <WIZ>
 R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.
 J. Biol. Chem. 248, 1631-1633, 1973
 A:Title: The primary structure of human plasminogen. II. The histidine loop of human
 A:Reference number: A92125; MUID:73149248
 A:Contents: annotation: active site
 R:Groskopf, W.R.; Summaria, L.; Robbins, K.C.
 J. Biol. Chem. 244, 3590-3597, 1969
 A:Title: Studies on the active center of human plasmin. Partial amino acid sequence o
 A:Reference number: A92048; MUID:69234739
 A:Contents: annotation: active site
 R:Trexler, M.; Vail, Z.; Pathy, L.
 J. Biol. Chem. 257, 7401-7406, 1982
 A:Title: Structure of the omega-amino-carboxylic acid-binding sites of human plasminog
 A:Reference number: A92382; MUID:82213905
 A:Contents: annotation: omega-amino-carboxylic acid binding sites
 R:Vail, Z.; Pathy, L.
 J. Biol. Chem. 259, 13690-13694, 1984
 A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essent
 A:Reference number: A92458; MUID:85054794
 A:Contents: annotation: fibrin binding site; omega-amino-carboxylic acid binding site
 R:Gao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.
 J. Biol. Chem. 271, 29461-29467, 1996
 A:Title: Kring domains of human angiotatin. Characterization of the anti-prolifera
 A:Reference number: A58811; MUID:97067211
 A:Contents: annotation
 R:Iljnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.
 Biochemistry 37, 4699-4702, 1998
 A:Title: Generation of an angiotatin-like fragment from plasminogen by stromelysin-1
 A:Reference number: A58812; MUID:9548733
 A:Contents: annotation
 R:Tulinsky, A.; Mulichak, A.M.
 submitted to the Brookhaven Protein Data Bank, July 1991
 A:Reference number: A5141; PDB:1PK4
 A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
 R:Tulinsky, A.; Wu, T.P.
 submitted to the Brookhaven Protein Data Bank, July 1991
 A:Reference number: A51488; PDB:2PK4
 A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454
 R:Wu, T.P.; Tulinsky, A.

submitted to the Brookhaven Protein Data Bank, August 1993
 A:Reference number: A51911; PDB:IPKR
 A:Contents: annotation: X-ray crystallography, 2.48 angstroms, residues 102-181
 R:Padmanabhan, K.; Tulinsky, A.
 Submitted to the Brookhaven Protein Data Bank, April 1994
 A:Reference number: A52408; PDB:IPMK
 A:Contents: annotation: X-ray crystallography, 2.25 angstroms, residues 377-454
 R:Tulinsky, A.; Mathews, I.I.
 submitted to the Brookhaven Protein Data Bank, December 1995
 A:Reference number: A65244; PDB:ICEA
 A:Contents: annotation: X-ray crystallography, 2.1 angstroms, residues 102-181
 R:Tulinsky, A.; Mathews, I.I.
 submitted to the Brookhaven Protein Data Bank, December 1995
 A:Reference number: A65245; PDB:ICEB
 A:Contents: annotation: X-ray crystallography, 2.1 angstroms, residues 102-181
 R:Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.
 Biochemistry 30, 10576-10588, 1991
 A:Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å
 A:Contents: annotation
 A:Reference number: A58819; PMID:92031502
 A:Contents: annotation
 R:Mu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.
 Biochemistry 30, 10589-10594, 1991
 A:Title: The refined structure of the epsilon-aminocaproic acid complex of human plasmin
 A:Contents: annotation
 A:Reference number: A58818; PMID:92031503
 A:Contents: annotation
 R:de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M.
 Biochemistry 31, 270-279, 1992
 A:Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2 Å
 A:Reference number: A39483; PMID:92118803
 A:Contents: annotation: X-ray crystallography, 2.4 angstroms
 R:Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.
 submitted to the Brookhaven Protein Data Bank, June 1995
 A:Reference number: A65980; PDB:IKRN
 A:Contents: annotation: X-ray crystallography, 1.67 angstroms, residues 376-454
 R:Rejante, M.; Llinas, M.
 submitted to the Brookhaven Protein Data Bank, August 1996
 A:Reference number: A65803; PDB:IHPJ
 A:Contents: annotation: conformation by (1)H-NMR, residues 103-181
 R:Rejante, M.; Llinas, M.
 submitted to the Brookhaven Protein Data Bank, August 1996
 A:Reference number: A65804; PDB:IHPK
 A:Contents: annotation: conformation by (1)H-NMR, residues 103-181
 R:Rejante, M.R.; Llinas, M.
 Eur. J. Biochem. 221, 927-937, 1994
 A:Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
 A:Reference number: S43645; PMID:94237157
 A:Contents: annotation: conformation by (1)H-NMR, residues 96-184
 R:Rejante, M.R.; Llinas, M.
 Eur. J. Biochem. 221, 939-949, 1994
 A:Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen
 A:Reference number: A58817; PMID:94237158
 A:Contents: annotation: conformation by (1)H-NMR
 C:Comment: Plasminogen is synthesized by the kidney and is present in plasma and many of
 C:Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKHU and
 C:Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITTHA2) immediately after
 C:Comment: Microplasmin is formed by autolytic cleavage of plasmin under artificial conditions
 C:Comment: Streptolysin 1 (see PIR:KCHU1) acts on plasminogen to produce angiotensin. The
 C:Comment: Streptolysin 1 (see PIR:KCHU1) acts on plasminogen to produce angiotensin. The
 C:Genetics:
 A:Gene: GDB:PLG
 A:Cross-references: GDB:119498; OMIM:173350
 A:Map position: 6q26-6q27
 A:Introns: 11/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 529/2
 C:Function:
 A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety
 A:Pathway: fibrinolysis
 C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
 C:Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydrolysis
 F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
 F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-810/Product: plasminogen #status experimental <PRO>
 F:20-96/Domain: activation peptide #status experimental <APT>
 F:79-46/Product: angiotensin #status experimental <AST>
 F:97-580,581-810/Product: plasmin #status experimental <MAT>
 F:97-580/Domain: plasmin chain A #status experimental <CHA>
 F:103-181/Domain: kringle homology <KR1>
 F:185-262/Domain: kringle homology <KR2>
 F:275-352/Domain: kringle homology <KR3>
 F:377-454/Domain: kringle homology <KR4>
 F:481-560/Domain: kringle homology <KR5>
 F:530-580,581-810/Product: microplasmin #status experimental <MPT>
 Query Match 33.2% Score 1371; DB 1; Length 810;
 Best Local Similarity 34.4%; Pred. No. 1,4e-88;
 Matches 292; Conservative 118; Mismatches 257; Indels 182; Gaps 22;
 QY 12 LGHVLHLLPLAIPIAEGQRKRRTIHERKSAKTLIKIDPAIKTKK--VNTAD 68
 DB 1 MEHEVLLLL-LFLKSGGE-----PLDDYVNTQGSLSFSV-----TKQLGAGSIE 47
 QY 69 QCANRCRRNGLPRTCKAFVDFDKARKOCIMPPFMSNGVKKERGFEDLEYENKDYLRNC 128
 DB 48 ECARKCEDE--EFTCAFYHNSKEQOCVMAENRKSIIIRM--RDVLEKKVYLSEC 103
 QY 129 IIGKRSYKGVSTITKSGIKQCPWSSMIPIHESPLSPSSRGKDLQENYCRNPRGEGGPW 188
 DB 104 KTGNGKRYRGMSMTKNGITCQKWSSTSPHRRSPATHTHSEGLENYCRNPDNDPGPW 163
 QY 189 CFTSNPVRREVCIDIPQCEVEYECMTNGESTYRGIMDHTSGKITQRMDOHPHRAKLP 248
 DB 164 CYTDPPEKRYDCIILCEE--ECMHCSGENDGKISTMSGLEQAWDSQSPHAGYIPS 222
 QY 249 RYPRKGFDDNCRNPDGCRPWCYTLDPRHMEVCAIKTCQDNTMNDTPLE--TTECI 306
 DB 223 KFPKNLKKKRYCRNPDDELAPMCFTTDPNKRWELCDIPRCT-----TPPPSGPTQCL 276
 QY 307 QGQEGYRGTYNTIWNIGIPCRQWDSQYRPHEDMTPEHFKCKDLRENYCRNPDGSESPCF 366
 DB 277 KGTGENYRGNAVAVVSGHTQGMASQTPPHNRRPEHNPCKNDENYCRNPDGKRAVWCH 336
 QY 367 TTDPNIVGCSQIPLNCDMSH-----GDDCYRGNGKNTKNGNISQTRSG 409
 DB 337 TTNSQVMEYC-KIPSCDSSPVSTEO LAPAPLPELVVDQCYHGDSQYRGTSSTTTTG 395
 QY 410 LTGSMV-----DK----- 417
 DB 396 KKGQSSWSPRHHQKRPENTPNAGLTMNCRNPDADKGPWCFTTDPVMEYCNLKKCS 455
 QY 418 -----NME-----DLHRHF 427
 DB 456 GTEASVAPRPVLLPVRVETPSEEDCMFNGCKYRGKRAITVTGTPCODMAAOEPRHHSI 515
 QY 428 WEPDASK--LNENYCRNPDDAHGPCYTGPNLLIPWDYCPISCEGDTPTIYNDHPV 484
 DB 516 FTPEPNRAGLEKKNYCRNPDGVCPCYTTPNPKLIDYCDVCAA-----PS 564
 QY 485 ISCAKTK-----QLRVNGIPTRTNI-GMVSILRYR-NKHIGGSLIKESWLTARQCF 536
 DB 565 FDGKPFVEPRKPCRGVAVGGCVAAHPHSPWQVSLTRFGHMFCCGTLISPEWVLTAAHCL 624
 QY 537 -----PSRDLKDYEWMLGIHDVHGRDEKCKQVLYNSQIYVGPESGLVLMKLARPAVLD 591
 DB 625 EKSPRPS-----SYKVLIGAHQEVNL--EPHQEILEVRLPLEPRKIDALLSSPAVIT 678
 QY 679 DKVAPALPSPNYVAVRTECFITGMEETQGTGAGLKEQLVLENNKVCNRYEFLNGR 738
 QY 652 LNSEIQTGAEMKIGSGPCEDGYGGLVCEQKMMVGLVIVPGGCAIPNPGIFGVVAY 711
 DB 739 VQSEILAGHLGAGTSCQDGSGLPVCFEKDYITLOGVTSWGLGCAPNKPGVAVRSR 798
 QY 712 YAKWIKRII 720

Db 799 FVTWIEGVM 807

RESULT 13

B30848
C:Accession: EC 3.4.21.7 precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 22-Jun-1999
R:Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5937-5965, 1989
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A:Reference number: A32869; MUID:89174660
A:Accession: B32869
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-810 <TOM>
A:Cross-references: GB:J04697; NID:g342272; PIDN:AAA36901.1; PID:g342273
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-9/Domain: plasminogen-related protein precursor homology <PLPH>
F:103-181/Domain: signal sequence #status predicted <SIG>
F:185-262/Domain: kringle homology <KR1>
F:275-352/Domain: kringle homology <KR2>
F:377-454/Domain: kringle homology <KR3>
F:481-560/Domain: kringle homology <KR4>
F:581-803/Domain: trypsin homology <TRY>
F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,324-342,665,760/Active site: His, Asp, Ser #status predicted

Query Match 33.2%; Score 1369; DB 2; Length 810;
Best Local Similarity 33.8%; Pred. No. 2e-88;

Matches 286; Conservative 127; Mismatches 257; Indels 176; Gaps 21;

Db 12 LOHVLHLLLPALPYAEGORRRNTIHEFKSAKTLIKIDPALKIKTKR--VMTAD 68
1 MEHKEVYLLLL-LFKSGGE-----PLDYYVNTKGASLSI-----TKKGAGASIE 47
Db 69 QCANRCRANKGLPTCAFYFDKARKKCLMFPPNSMSSGVKKEFGHEFDLYENKDYTRNC 128
48 ECARKEEEL--EFTCSFYHSKEDOCVIAENKRSIVFRM--RDYVLFEEKVYVLESEC 103
Db 129 IICKSGSYKTSVITKSGICOPWMSWIRPHEHSFLPSYSGKDIOENYCNPRREEGPW 188
104 KTGNGKNYRGTMTKRTGICQKMSISPHRPTFSFATHPSEGLEENYCNRPDNDGGGPW 163
189 CFTSNPEVRYEVDIPQSEVECMTCNGESYRGLMDHTESGICQRMWDQTPHNRKFLPE 248
164 CYTTDEEREDYCDIPEC--EDECMSGSENYDKISKTMSGLEGCQAMDQSPNHNHYTPS 222
Db 249 RYDDKGFDDNYCNPDPGPRPWCYTLDPHRMWECAIKTADNTMNDTVPLE--TTECI 306
223 KFPNKLKKNYCNPRGEPFPCFTDPNKRWELCDIPRT-----TPPPSSGPTYOQL 276
Db 307 QOGGEGYKGTIVTNGIIPQRMDSQYRPHEDMTPEFKKDLRENYCRPDGSESPWCF 366
277 KGTGENYRGDAVATVSGHCHGWSAQTPIHNTPEPFCKNIDENYCRNPDEKAPWCY 336
Db 367 TTDPRNRYGCSQIPNCDMSH-----GDCYRGNKNYMGNLISQTRSG 409
337 TNSOVRWEYC--KIPSCSSPVSTELDPTAPELTPVQECYHGGOSYRGSSTTTTG 395
Db 410 LFCSMYDKMIEDLHR-----HIFWE----- 429
396 KKCQSSMSWTPIHHEKTPENFPNAGLTMNYCRNPADKGRPWCFTTDPVSWEYCNLKKS 455
Db 430 -----PDA----- 432
456 GTGSAVAPPAVLAQDAPATPSEEDCMFNGCKYRGKAKATYVTGTCQEMAQDEPHSHRI 515

Db 433 -----SKLNENYCRNPDDDAHGPWCYTGNPILPWCYSPISRCGDDTPTIVLMDHPV 484
516 FTEPTNPRAGLEKNYCRNPDDGVDGPMCYTTPNPKLFYCDVPOCAASS----- 564
Db 485 ISCAKTK-----QLRYVNG--IPRTNIGMWVSLRYR--NKHCGSLKESVLTAROCF 536
555 FDCCGPOVEPKKCPGRVVGCGVAPHSMPWQISLTRILGMHFCGDTLISPMVLTAAACL 624
Db 537 PSRLDKL-TEAMLGIH-DVHGKDEKCKQVNLVSGVPGSDVLMKRLARPVLDFEV 594
625 EKSSRPSFYKVILGAHREHV--LEPHQDELEVSKMSEPARADIALKLSSPAIITDKV 681
Db 595 STIDLPNKGCIIPKETSQSVYGMWYTGILNYDGLIRVAHLIYMGNEKCSOHNRKQVTLNE 654
682 IPACLPSPNYVADRTCEFTIGMETQGTGAGLKEARLFLVINKYCNRYEFLNGTYKT 741
Db 655 SETCAGAEKIGSGPOEGYGGPLVCEQHKRMVGLVTPRGCAIIPRRCIFVAVATYAK 714
742 TELCAGHLAGTIDSCQDSDGSPVLCFEKDKYILLOGVTSWGLGCAKAPMKPGVYVAVSRFVT 801
Db 715 WIKRII 720
Db 802 WIEGVM 807

RESULT 14

146260
C:Accession: EC 3.4.21.7 precursor - western European hedgehog
C:Species: Erinaceus europaeus (western European hedgehog)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
R:Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; F
J. Biol. Chem. 270, 24004-24009, 1995
A:Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipop
A:Reference number: I46259; MUID:96025778
A:Accession: I46260
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-810 <LAW>
A:Cross-references: EMBL:033171; NID:g1046360; PID:g1046361
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C:Keywords: hydrolase; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:379-456/Domain: kringle homology <KR4>
F:482-561/Domain: kringle homology <KR5>
F:582-803/Domain: trypsin homology <TRY>

Query Match 32.8%; Score 1353.5; DB 2; Length 810;
Best Local Similarity 35.1%; Pred. No. 2.4e-87;

Matches 277; Conservative 110; Mismatches 255; Indels 147; Gaps 16;

Db 61 TKR-----VNTADCANCTRNKGLPFTCAFYFDKARKKCLMFPPNSMSSGVKKEFGHEHD 117
37 TKQOLSVGSTECAVVC--EKETSFICRSFOYHSKEDOCVIAENKRSIVFLRM--RDYI 92
Db 118 LYENKDYIRNCTIGKGRSYKTSVITKSGICQPMSSMIRPHEHSFLPSYSGKDIOENYCN 177
93 LFEKMYLYSECKVGNKGYRGTYSKRTGLTQKWSAELEPHRPFREPDENPSEGLDQNYC 152
Db 178 RNRGEGGPWCFTSNPEVRYEYCDIPQSEVECMTCNGESYRGLMDHTESGICQRMWDH 237
153 RNPNDPKGPWCYTMDPEVRYEYCEIQC--EDECMSGQANYVGIKISPTMSGLECOQWDS 211
Db 238 QPFRHKKFLPERYPCGPDNCRNPDPGPRPWCYTLDPHRMWECAIKTADNTMNDTD 297
212 QTPHPHGFIPSKFPSKNLKNYCNRPDGEPRPWCFTMDNKRKWECDIPRT-----TP 265
Db 298 VPLE--TTECIOGQGGYRGVNTIWNIGIPCORWDSQYRPHEDMTPEFKKDLRENYCR 355

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2002, 17:46:19 ; Search time 13.55 Seconds

(without alignments)
2080.283 Million cell updates/sec

Title: US-09-423-516-2

4126

Sequence: 1 MMYTKLLPALLQHVLLHL.....VAYYAKNIKILITRYKVPQS 728

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

105224 seqs, 38719550 residues

Number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4126	100.0	728	1 HGF_HUMAN	P14210 homo sapien
2	3804	92.2	728	1 HGF_MOUSE	Q08048 mus musculu
3	3802	92.1	728	1 HGF_RAT	P17945 rattus norv
4	1652	40.0	711	1 HGFL_HUMAN	P26927 homo sapien
5	1602	38.8	716	1 HGFL_MOUSE	P26928 mus musculu
6	1446.5	35.1	790	1 PLMN_PIG	P06867 sus scrofa
7	1408.5	34.1	812	1 PLMN_BOVIN	P06868 bos taurus
8	1379.5	33.4	810	1 PLMN_MOUSE	P20918 mus musculu
9	1370	33.2	810	1 PLMN_HUMAN	P00747 homo sapien
10	1369	33.2	810	1 PLMN_MACMU	P12545 macaca mula
11	1353.5	32.8	810	1 PLMN_ERIEU	Q29485 erinaceus e
12	1212.5	29.4	4548	1 APOA_HUMAN	P08319 homo sapien
13	1188.5	28.8	1420	1 APOA_MACMU	P14417 macaca mula
14	583	14.1	343	1 PLMN_SHEEP	P81486 ovis aries
15	559.5	13.6	559	1 TPA_RAT	P19637 rattus norv
16	552.5	13.4	338	1 PLMN_HORSE	P80010 equus cabal
17	551.5	13.4	559	1 TPA_MOUSE	P11214 mus musculu
18	548.5	13.3	333	1 PLMN_CANFA	P80009 canis fami
19	543	13.2	625	1 THRB_BOVIN	P00735 bos taurus
20	539.5	13.1	566	1 TPA_BOVIN	Q028198 bos taurus
21	532.5	12.9	562	1 TPA_HUMAN	P00750 homo sapien
22	531.5	12.9	618	1 THRB_MOUSE	P19921 mus musculu
23	527.5	12.8	622	1 THRB_HUMAN	P00734 homo sapien
24	518	12.6	617	1 THRB_RAT	P18292 rattus norv
25	510.5	12.4	603	1 FA12_CAVPO	Q04962 cavia porce
26	497.5	12.1	653	1 HGFA_MOUSE	Q04962 mus musculu
27	480.5	11.6	655	1 HGFA_HUMAN	Q04756 homo sapien
28	449	10.9	325	1 PLMN_PETMA	P33574 petromyzon
29	449	10.9	615	1 FA12_HUMAN	P00748 homo sapien
30	426.5	10.3	593	1 FA12_BOVIN	P98140 bos taurus
31	393	9.5	433	1 UROK_BOVIN	Q05589 bos taurus
32	350	9.5	431	1 UROK_HUMAN	P00749 homo sapien
33	388	9.4	433	1 UROK_MOUSE	P06869 mus musculu

34	386	9.4	442	1 UROK_PIG	P04185 sus scrofa
35	385	9.3	434	1 UROK_CHICK	P15120 gallus gall
36	382.5	9.3	432	1 UROK_RAT	P29598 rattus norv
37	382.5	9.3	433	1 UROK_PAPCY	P16227 papio cynoc
38	375	9.1	1035	1 ENTR_BOVIN	P98072 bos taurus
39	373.5	9.1	477	1 URT2_DESRO	P15638 desmodus ro
40	373	9.0	477	1 URT1_DESRO	P98119 desmodus ro
41	372	9.0	273	1 MCT7_MOUSE	Q02844 mus musculu
42	369	8.9	431	1 URT6_DESRO	P98121 desmodus ro
43	358	8.7	273	1 MCT7_RAT	P27435 rattus norv
44	358	8.7	394	1 URT6_DESRO	P49150 desmodus ro
45	357	8.7	1069	1 ENTR_MOUSE	P97435 mus musculu

ALIGNMENTS

RESULT 1
HGF_HUMAN STANDARD: PRT; 728 AA.
ID HGF_HUMAN
AC P14210; O99BXL9;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
DE (Hepatopoietin A).
CN HGF OR HPTA.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91340155; PubMed=1831432;
RA Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;
RT "Organization of the human hepatocyte growth factor-encoding gene.";
RN Gene 102:213-219(1991).
[2]
RP SEQUENCE FROM N.A.
RX TISSUE=Placenta;
RA MEDLINE=89392017; PubMed=2528952;
RA Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,
RA Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,
RA Gohda E., Daikuhara Y., Kitamura N.;
RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte growth factor.";
RN Biochem. Biophys. Res. Commun. 163:967-973(1989).
[3]
RP SEQUENCE FROM N.A.
RX TISSUE=leukocyte;
RA MEDLINE=91025062; PubMed=2145836;
RA Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T.,
RA Asami O., Hagiya M., Nakamura T., Shimizu S.;
RT "Isolation and expression of cDNA for different forms of hepatocyte growth factor from human leukocyte.";
RN Biochem. Biophys. Res. Commun. 172:321-327(1990).
[4]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 55-73 AND 495-520.
RX TISSUE=liver;
RA MEDLINE=90066676; PubMed=2531289;
RA Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,
RA Sugimura A., Tashiro K., Shimizu S.;
RT "Molecular cloning and expression of human hepatocyte growth factor.";
RN Nature 342:440-443(1989).
[5]
RP SEQUENCE FROM N.A.
RX TISSUE=Embryonic fibroblast;
RA MEDLINE=9134393; PubMed=1831266;
RA Weidner K.M., Arakaki N., Hartmann G., Vandekerckhove J., Weingart S.,
RA Riederer H., Fomatsch C., Tsubouchi H., Hishida T., Daikuhara Y.,
RA Birchemer W.;
RT "Evidence for the identity of human scatter factor and human hepatocyte growth factor.";

RL Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Courtney L., Elliot G., Angell S.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 249-695 FROM N.A.
 RA MEDLINE=91369928; PubMed=1832556;
 RT Miyazawa K., Kitamura A., Kitamura N.;
 RT "Structural organization and the transcription initiation site of the
 human hepatocyte growth factor gene.";
 RT Biochemistry 30:9170-9176(1991).
 RN [8]
 RP SIGNAL SEQUENCE CLEAVAGE SITE.
 RA MEDLINE=91207365; PubMed=1826837;
 RA Yoshiyama Y., Areakaki N., Naka D., Takahashi K., Hirano S., Kondo J.,
 RA Nakayama H., Gohda E., Kitamura N., Tsubouchi H., Ishii T.,
 RA Hishida T., Dokuhara Y.;
 RT "Identification of the N-terminal residue of the heavy chain of both
 native and recombinant human hepatocyte growth factor.";
 RT Biochem. Biophys. Res. Commun. 175:660-667(1991).
 RN [9]
 RP CARBOHYDRATE-LINKAGE SITE 476.
 RA MEDLINE=93129192; PubMed=1482348.
 RA Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,
 RA Nakamura T., Shimizu S.;
 RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
 on the alpha chain.";
 RT Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
 RN [10]
 RP MUTAGENESIS.
 RA MEDLINE=92331602; PubMed=1321034;
 RA Lokker J.B., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
 RA Baker J.B., Godowski P.J.;
 RT "Structure-function analysis of hepatocyte growth factor:
 RT Identification of variants that lack mitogenic activity yet retain
 high affinity receptor binding.";
 RT EMBO J. 11:2503-2510(1992).
 RN [11]
 RP STRUCTURE BY NMR OF 31-127.
 RA MEDLINE=98154323; PubMed=9493272;
 RA Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,
 RA Rubin J.S., Bottaro D.P., Byrd R.A.;
 RT "The solution structure of the N-terminal domain of hepatocyte growth
 factor reveals a potential heparin-binding site.";
 RT Structure 6:109-116(1998).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.
 RA MEDLINE=99036858; PubMed=9817840;
 RA Ullrich M., Lokker N.A., Godowski P.J., de Vos A.M.;
 RT "Crystal structure of the NK1 fragment of human hepatocyte growth
 factor at 2.0-A resolution.";
 RT Structure 6:1383-1393(1998).
 CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
 CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS
 CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
 CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D90334; BAA14348.1; -
 CC EMBL: D90318; BAA14348.1; JOINED.
 DR EMBL: D90319; BAA14348.1; JOINED.
 DR EMBL: D90320; BAA14348.1; JOINED.
 DR EMBL: D90322; BAA14348.1; JOINED.
 DR EMBL: D90323; BAA14348.1; JOINED.
 DR EMBL: D90324; BAA14348.1; JOINED.
 DR EMBL: D90325; BAA14348.1; JOINED.
 DR EMBL: D90326; BAA14348.1; JOINED.
 DR EMBL: D90327; BAA14348.1; JOINED.
 DR EMBL: D90328; BAA14348.1; JOINED.
 DR EMBL: D90329; BAA14348.1; JOINED.
 DR EMBL: D90330; BAA14348.1; JOINED.
 DR EMBL: D90331; BAA14348.1; JOINED.
 DR EMBL: D90332; BAA14348.1; JOINED.
 DR EMBL: D90333; BAA14348.1; JOINED.
 DR EMBL: M29145; AAA52650.1; -
 DR EMBL: M60718; AAA52648.1; -
 DR EMBL: X16323; CAA34387.1; -
 DR EMBL: M73239; AAA64239.1; -
 DR EMBL: M73240; AAA64297.1; -
 DR EMBL: AC004960; AAC71655.1; -
 DR EMBL: M75983; AAG53460.1; -
 DR EMBL: M75972; AAG53460.1; JOINED.
 DR EMBL: M75973; AAG53460.1; JOINED.
 DR EMBL: M75974; AAG53460.1; JOINED.
 DR EMBL: M75975; AAG53460.1; JOINED.
 DR EMBL: M75976; AAG53460.1; JOINED.
 DR EMBL: M75977; AAG53460.1; JOINED.
 DR EMBL: M75978; AAG53460.1; JOINED.
 DR EMBL: M75979; AAG53460.1; JOINED.
 DR EMBL: M75980; AAG53460.1; JOINED.
 DR EMBL: M75981; AAG53460.1; JOINED.
 DR EMBL: M75982; AAG53460.1; JOINED.
 DR PIR: JH0579; JH0579.
 DR PIR: S06794; S06794.
 DR PDB: 2HGF; 24-JUN-98.
 DR PDB: 1BHT; 18-NOV-98.
 DR MEROPS; S01.976; -
 DR GlycoSiteDB; P14210; -
 DR MIM; 142409; -
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00051; kringle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS; PRO0722; CHYMOTRYPSIN.
 DR PRINTS; PRO0018; KRINGLE.
 DR SMART; SM00130; KR; 4.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00020; TRYF_SPC; 1.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS0070; KRINGLE_2; 4.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR Growth factor; Kringle; Glycoprotein; serine protease homolog;
 KW Signal; 3D-structure.
 FT SIGNAL 1 31
 FT CHAIN 32 494
 FT CHAIN 495 728
 FT MOD_RES 32 32
 FT DOMAIN 32 127
 FT DOMAIN 128 206
 FT DOMAIN 211 288
 FT DOMAIN 305 383
 FT DOMAIN 391 469
 FT DOMAIN 495 728
 FT DISULFID 70 96
 FT DISULFID 74 84
 FT DISULFID 128 206
 FT DISULFID 149 189
 FT DISULFID 177 201
 FT DISULFID 487 604
 INTERCHAIN (BY SIMILARITY).

Query Match 100.0%; Score 4126; DB 1; Length 728;
 Best Local Similarity 100.0%; Pred. No. 6.5e-301;
 Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWVTKLLPALLLOHVLHLLHLLPIAIPVAGGKRRRNTTHEFKSAKTLIKIDPALKIK 60
 DB 1 MWVTKLLPALLLOHVLHLLHLLPIAIPVAGGKRRRNTTHEFKSAKTLIKIDPALKIK 60
 QY 61 TKVNTADOCANCTRNKGLPFTCKAFVDDKAKOCLMPFNMSGVKKEFHEEDLYE 120
 DB 61 TKVNTADOCANCTRNKGLPFTCKAFVDDKAKOCLMPFNMSGVKKEFHEEDLYE 120
 QY 61 TKVNTADOCANCTRNKGLPFTCKAFVDDKAKOCLMPFNMSGVKKEFHEEDLYE 120
 DB 61 TKVNTADOCANCTRNKGLPFTCKAFVDDKAKOCLMPFNMSGVKKEFHEEDLYE 120
 QY 121 NKDYINNCIIIGKRSYKGYSTIKSGIKCOPMSMI PHEHSFLPSSYRGKDLQENCRNP 180
 DB 121 NKDYINNCIIIGKRSYKGYSTIKSGIKCOPMSMI PHEHSFLPSSYRGKDLQENCRNP 180
 QY 181 RGESEGPWCFTSNPEVREYECVDPQSEVECMTCNGESYRGMLDHTESGKICQRMHDQTP 240
 DB 181 RGESEGPWCFTSNPEVREYECVDPQSEVECMTCNGESYRGMLDHTESGKICQRMHDQTP 240
 QY 241 HRKELPERYPOKGPNDNCRNPDGP RPWCYTLDPHTREYCAITTCADNTMNDVPL 300
 DB 241 HRKELPERYPOKGPNDNCRNPDGP RPWCYTLDPHTREYCAITTCADNTMNDVPL 300
 QY 301 ETTECTIOGGEGYRGVNTIMNGIPCORMDQY PHEHDMTPENFKCKDLRENCRNPDGS 360
 DB 301 ETTECTIOGGEGYRGVNTIMNGIPCORMDQY PHEHDMTPENFKCKDLRENCRNPDGS 360
 QY 361 ESPWCFETPDNIRVGYCSQIPNCDMSHGDCYRGNGKNYGNLSQTRSGLTCSMDKNME 420
 DB 361 ESPWCFETPDNIRVGYCSQIPNCDMSHGDCYRGNGKNYGNLSQTRSGLTCSMDKNME 420
 QY 421 DLARHIFEMERDASKLENNCRNPDDAHGWCCTGNPLIPMDYCPISRCGDTPTTVNL 480
 DB 421 DLARHIFEMERDASKLENNCRNPDDAHGWCCTGNPLIPMDYCPISRCGDTPTTVNL 480
 QY 481 DHPVISCATKQOLRVYNGIPTRTNIGMNVSLARYRNKHICGSLIKESWVLTARQCPSPRD 540
 DB 481 DHPVISCATKQOLRVYNGIPTRTNIGMNVSLARYRNKHICGSLIKESWVLTARQCPSPRD 540
 QY 541 LKDYEARMLGIDHVGSRGDECKQVNLVYSGESDLYMLKARPAVLDDFVSTIDLP 600
 DB 541 LKDYEARMLGIDHVGSRGDECKQVNLVYSGESDLYMLKARPAVLDDFVSTIDLP 600
 QY 601 NNGCTIPERTSCSVGMGTGLINVDGLRYAHLXTIMNGEKSQHHRGKVTLMSEICAG 660
 DB 601 NNGCTIPERTSCSVGMGTGLINVDGLRYAHLXTIMNGEKSQHHRGKVTLMSEICAG 660
 QY 661 AEKISGPGCEGIDGYPVCEQHMKRMVLAGVYVGRGCAIPNRPGIFRVAVYAKWIKHII 720
 DB 661 AEKISGPGCEGIDGYPVCEQHMKRMVLAGVYVGRGCAIPNRPGIFRVAVYAKWIKHII 720
 QY 721 LTKYKPOS 728
 DB 721 LTKYKPOS 728

RESULT 2
 HGF_MOUSE
 ID HGF_MOUSE STANDARD; PRT; 728 AA.
 AC 008048; 064007; 061662;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hepatocyte growth factor precursor (Scatter factor) (SF)
 DE (Hepatopoietin A).
 DE HGF.
 GN
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1)

RP SEQUENCE FROM N.A., AND SEQUENCE OF 496-504.
 RC TISSUE=Mammary fibroblast;
 RX MEDLINE=94183257; PubMed=8135822;
 RA Sasaki M., Nishio M., Sasaki T., Enami J.,
 RT "Identification of mouse mammary fibroblast-derived mammary growth
 RT factor as hepatocyte growth factor.";
 RL Biochem. Biophys. Res. Commun. 199;772-779(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94363381; PubMed=8081873;
 RA Lee C.C., Kozak C.A., Yamada K.M.;
 RT "Structure, genetic mapping, and expression of the mouse Hgf/scatter
 RT factor gene.";
 RL Cell Adhes. Commun. 1:101-111(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94060105; PubMed=8241272;
 RA Liu Y., Michalopoulos G.K., Zarnegar R.;
 RT Molecular cloning and characterization of cDNA encoding mouse
 RT hepatocyte growth factor.";
 RL Biochim. Biophys. Acta 1216:299-303(1993).
 CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
 CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS
 CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
 CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1- ALTERNATIVE PRODUCTS: A SHORT FORM OF HGF IS PRODUCED BY
 CC ALTERNATIVE RNA SPLICING. THE SEQUENCE SHOWN HERE IS THAT OF THE
 CC LONG FORM.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D10212; BAA01064.1; -;
 DR EMBL: D10213; BAA01065.1; -;
 DR EMBL: S71816; AAB31855.1; -;
 DR EMBL: X72307; CAA51054.1; ALT_INIT.
 DR HSSP: P14210; 1BHT.
 DR MGD: MGI:96079; Hgf.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; PAN_app.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00051; kringle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR SMART: SM00130; KR; 4.
 DR SMART: SM00473; PAN_AP; 1.
 DR SMART: SM00020; TRYP_SP; 1.
 DR PROSITE: PS00021; KRINGLE_1; 4.
 DR PROSITE: PS50070; KRINGLE_2; 4.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR Growth factor; Kringle; Glycoprotein; Serine protease homolog;
 KW Signal; Alternative splicing.
 FT SIGNAL 1 32
 FT CHAIN 33 495
 FT CHAIN 496 728
 FT MOD_RES 33 33
 FT (BY SIMILARITY).
 FT (BY SIMILARITY).

```

FT DOMAIN 33 128 PAB.
FT DOMAIN 129 207 KRINGLE 1.
FT DOMAIN 212 289 KRINGLE 2.
FT DOMAIN 306 384 KRINGLE 3.
FT DOMAIN 392 470 KRINGLE 4.
FT DOMAIN 496 728 SERINE PROTEASE-LIKE.
FT DISULFID 71 97 BY SIMILARITY.
FT DISULFID 75 85 INTERCHAIN (BY SIMILARITY).
FT DISULFID 488 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 656 656 MISSING (IN SHORT ISOFORM).
FT VARSPLIC 163 167 N -> K (IN REF. 2).
FT CONFLICT 344 344 V -> L (IN REF. 2).
FT CONFLICT 479 479 V -> H (IN REF. 3).
FT CONFLICT 564 564 R -> H (IN REF. 3).
SQ SEQUENCE 728 AA: 82944 MW: A0381FC497534328 CRC64;

```

Query Match 92.28; Score 3804; DB 1; Length 728;

Best Local Similarity 90.88; Pred. No. 7.5e-277; Matches 660; Conservative 39; Mismatches 26; Indels 2; Gaps 1;

```

QY 1 MWYKLLPALLQHVLLHLLLPYAPAEQORRRNTIHEFKSAKTTLIKIDPAKIK 60
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 MNGTKLLPVLLOHVLLHLLHVAIYAEQORRRNTIHEFKSAKTTLIKIDPAKIK 61
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 TKKYNVADQCANRCKTRKKGLEPTCKAFVFDKARKOCLEPPNNSSGYKKKEGHEFDLYE 120
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 TKKYNVADQCANRCKTRKKGLEPTCKAFVFDKARKOCLEPPNNSSGYKKKEGHEFDLYE 121
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 NKQVIRCIIGKGRSYGTYSITKSGIKCOPMSMIPIHEHFLSSYRGKLOJENCNP 180
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 122 NKQVIRCIIGKGRSYGTYSITKSGIKCOPMSMIPIHEHFLSSYRGKLOJENCNP 181
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 RGEGRGWCFTSNPEVRYEVDIPQCESEVECMTCNGESYRGKLOJENCNP 240
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 RGEGRGWCFTSNPEVRYEVDIPQCESEVECMTCNGESYRGKLOJENCNP 241
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 HRHKKFLPERIPDKFDNVCNPNPQGRPMCYTLDPTIRWETCAIKTCADNTMDTVPL 300
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 242 HRHKKFLPERIPDKFDNVCNPNPQGRPMCYTLDPTIRWETCAIKTCADNTMDTVPL 301
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 ETTPCIGOGGSGYGTNTINGIIPCORMDQYRHEHMPENKCKDLRENYCRNPDS 360
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 302 ETTPCIGOGGSGYGTNTINGIIPCORMDQYRHEHMPENKCKDLRENYCRNPDS 361
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 ESPWCFTTDPNIRVGYCSQIIPNCDMSHGQDCYRGNGKNYMGNLQSRTSGLTSMWDMKME 420
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 ESPWCFTTDPNIRVGYCSQIIPNCDMSHGQDCYRGNGKNYMGNLQSRTSGLTSMWDMKME 421
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 DLHRIHTEWPDASKLNEYCNPNDDAHGPRCTYGNPLIPMDYCPISCEEDTPTIYNL 480
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 422 DLHRIHTEWPDASKLNEYCNPNDDAHGPRCTYGNPLIPMDYCPISCEEDTPTIYNL 481
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 DHPVISCATKQQLVNVGIGIPRTINIGMWVSLRYRNKNIHCGSLIKESVYLAROCFSPR 539
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 482 DHPVISCATKQQLVNVGIGIPRTINIGMWVSLRYRNKNIHCGSLIKESVYLAROCFSPR 541
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 540 -DLKDYBAMLGIDHVGSRGDEKCKQVNLVNSQLVYGPESGSDLVLLKLARPAVLDDFVSTID 598
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 542 KDLKDYBAMLGIDHVGSRGDEKCKQVNLVNSQLVYGPESGSDLVLLKLARPAVLDDFVSTID 601
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 599 LPNKGCTIPREKTSYGVGWTGLINTDGLLVAHLIYMGKEKSSQHHKRGVTLNESITC 658
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 602 LPNKGCTIPREKTSYGVGWTGLINTDGLLVAHLIYMGKEKSSQHHKRGVTLNESITC 661
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 659 AGAKKISGPEEGYGGGLVCEGKHMVNLVIVPGRGCAIPNPGIVVRAVYAAKWIHK 718
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 662 AGAKKISGPEEGYGGGLVCEGKHMVNLVIVPGRGCAIPNPGIVVRAVYAAKWIHK 721
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 719 IILTYKV 725

```

```

DB 722 VILTYKL 728
  :|||||:
RESULT 3
HGF_RAT ID HGF_RAT STANDARD; PRT; 728 AA.
AC P17945;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
DE (Hepatopoietin A).
GN HGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=90222197; PubMed=2139229;
RA Toshito K., Hagiya M., Nishizawa T., Seki T., Shimonishi M.,
RA Shimizu S., Nakamura T.;
RT "Reduced primary structure of rat hepatocyte growth factor and
  expression of the mRNA in rat tissues.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=91031482; PubMed=2146117;
RA Okajima A., Miyazawa K., Kitamura N.;
RT "Primary structure of rat hepatocyte growth factor and induction of
  its mRNA during liver regeneration following hepatic injury.";
RL Eur. J. Biochem. 193:375-381(1990).
CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATRE PARENCHYMAL
CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS
CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D90102; BAA1413.1; -
DR EMBL: X54400; CAA38266.1; -
DR PIR: S13211; S13211.
DR PIR: A35644; A35644.
DR HSSP: P14210; 1BHT.
DR MEROPS: S01.978; -
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Trypsin.
DR Pfam: PF000051; kringle_4.
DR Pfam: PF000024; PAN_1.
DR Pfam: PF000089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR SMART: SM00130; KR: 4.
DR SMART: SM00473; PAN_Ap_1.
DR SMART: SM00020; TRYP_SEC; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.

```



```

DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp-Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
KW Kringle; Glycoprotein; Serine protease homolog; signal;
KW Polymorphism.
FT SIGNAL 1 18
FT CHAIN 19 711
FT DOMAIN 32 109
FT DOMAIN 110 186
FT DOMAIN 191 268
FT DOMAIN 283 361
FT DOMAIN 370 448
FT DOMAIN 484 711
FT DISULFID 56 78
FT DISULFID 60 66
FT DISULFID 110 186
FT DISULFID 131 169
FT DISULFID 157 181
FT DISULFID 191 268
FT DISULFID 194 324
FT DISULFID 212 251
FT DISULFID 240 263
FT DISULFID 283 361
FT DISULFID 304 343
FT DISULFID 332 355
FT DISULFID 370 448
FT DISULFID 391 431
FT DISULFID 419 443
FT DISULFID 468 588
FT DISULFID 507 523
FT DISULFID 602 667
FT DISULFID 632 646
FT DISULFID 657 685
FT CARBOHYD 72 72
FT CARBOHYD 296 296
FT CARBOHYD 615 615
FT VARIANT 13 13
FT VARIANT 13 13
FT VARIANT 212 212
FT VARIANT 212 212
FT CONFLICT 623 623
FT CONFLICT 711 AA: 80379 MM: 596ED21F180290E4 CRC64:
SQ SEQUENCE 711 AA: 80379 MM: 596ED21F180290E4 CRC64:

Query Match 40.0%; Score 1652; DB 1; Length 711;
Best Local Similarity 43.8%; Pred. No. 5,2e-116;
Matches 317; Conservative 103; Mismatches 265; Indels 38; Gaps 14;

```

```

QY 314 RCTVNTANGIPCOQRWDSQVPEHDMTPENFKCKDLRENYCNRPDGSESFPCFTTDPNIR 373
DB 292 RGTANTTAGVYCOQRWDAIPHOHRTPEKYACKDLRENYCNRPDGSEAPMCFTLRPGNR 351
QY 374 VGYCSQIPNC-DMSHGQDCYRGNGKNYMGNSQTRSGLTCSMMDKNMEDLHRIHWEPPDA 432
DB 352 AAFVQYIRRCTDVPDRPODCYHGAEGYRGCTVSKTRGVCCQRMASAEPTPKPQFTFTSEPH 411
QY 433 SKLTENYCCNRPDGDAHGPCYTGNGNPLIPNDYCPISGCEBDTPTIYNLDHP-----VISCA 488
DB 412 AOLENEPCNRPDGSDSHGPCYTYMDPPTPDYCALRCADDPSP1--LDPPQVOYFEKKG 469
QY 489 K-----TKOLRVYNGIPRTNIGMWSLRVR-NKHICGSGSLKESVTLAROCFSP- 538
DB 470 KAVDRLDQKRSRLRYVGGP--GNSPMYVSLNRQGHCGGSLYKEOMITLAROCFSSC 527
QY 539 -RDLDYEAMLGIDHYHGRDEKQOVLYNSQLYVPEGSDYLVLKLARPAVLDDFVSTI 597
DB 528 HMLPGYEWLGLTPONPQHGEPSSLQRPVAKMCGSPSGSLVLKLEKSVTLNQRVALI 587
QY 598 DLPNNGCTIPEKTSQSVYWGVTGLINQGLRVHLYTMGKCSQHHRGKVTLNESFI 657
DB 588 CLPPEMYVVPQTKCEIAGMGETKGTGNDTILNALVINSQECNIKIRGRV--RESEM 645
QY 658 CAGAEKIGSGPECGDYGCEHMKRMVLYGVPRGCAIPNRPGRIPVAVYAKMTH 717
DB 646 CTEGLLAPVAGNCEGDIYGGFLACFTINCWLVBSIIIPNRVCAKSRMPAVTRVSVVDIMH 705
QY 718 KII 720
DB 706 KVM 708

```

```

RESULT 5
HGFL_MOUSE
ID HGFL_MOUSE STANDARD: PRT; 716 AA.
AC P26928;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hepatocyte growth factor-like protein precursor (Macrophage
stimulatory protein) (MSP).
GN MST1 OR HGFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Liver;
RX MEDLINE=92002017; PubMed=1832957;
RA Fritzenner Degen S.J., Stuart L.A., Han S., Jamison C.S.;
RT "Characterization of the mouse cDNA and gene coding for a hepatocyte
growth factor-like protein: expression during development.";
RL Biochemistry 30:9781-9791(1991).
CC -!- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA
CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT
CONSERVED.
CC -!- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND
ADRENAL.
CC -!- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION.
CC JUST BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS
CC STABLE AFTERWARDS.
CC -!- PM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
CC HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
CC POLYPEPTIDES.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


Best Local Similarity: 36.9%; Pred. No. 1,3e-100; Matches 294; Conservative 114; Mismatches 232; Indels 157; Gaps 20

FT	DISULFID	701	768	BY SIMILARITY.
FT	DISULFID	731	747	BY SIMILARITY.
FT	DISULFID	758	786	BY SIMILARITY.
SO	SEQUENCE	812 AA;	90846 MW;	D34A744AC2256F8 CRC64

Query Match	33.4%	Score 1379.5;	DB 1;	length 812;
Best Local Similarity	34.4%	Pred. No. 1.4e-95;		
Matches 288;	Conservative 119;	Mismatches 262;	Indels 167;	Gaps 22

QY	15	VILHILLLPLAIPAEOSORRRNRIHEFFKSAAKTLIKIDPALKIKRKVNTA----	DOCA	71
Db	7	ILFLFULL-----KPO-----GSLDSGTISTOGASLFL-----	TKOLAAGVSDCL	50
QY	72	NCRTRNKGLPFCKAEVDFARAKQCIAMPENSMSSGVKKEFGEFDLYENKDYIRNCTIG		131
Db	51	AKCEGE--TDFVCSFQYHSKEQOCVIAENSKTSIIIR--RDVILFEKRYVLECKTG		106
QY	132	KGRYKKTVAITITSGICQPMSSMIIPHEHSFLPSYKGLDQENYCNPNRGEEGPFCYT		191
Db	107	IGNMYRQTMSTRKTSGVACQKMGATFPPVPMVPSPTIHNEGLEEYCNPNPNDDEGPFCYT		166
QY	192	SNEPAREVDCIPDCSEVECMTCNGESYRGLMDHETEGKICQMRDQDTPRHAKLIPREY		251
Db	167	TDPDKRDYCNIPCEE--ECMYCSGEYEBKISKTMSGLDQAMDSPPAHGIIIPKFP		225
QY	252	DKGFDNYCNRPDQOPRPWCYTLDPHTRWECALIKTADNTMDTDVPLETTECIGQGE		311
Db	226	SKNLKNMYCNHPDDEPRPWCFTTDPTRKMEWCIDIPRCT-----PPPPSPFYOCCLKRGE		281
QY	312	GYRGTVATINGIIPCOAWDQYRHEHMTPENFCKDLRENYCNPNRGSPPMCFITDPN		371
Db	282	NYRGTAVTVSGTKCQMSDEQTPRHNRTPENPCKMLEEYCNPNPDEGATPAWCYTTDSQ		341
QY	372	IRVQCSQIINCDSMSH-----QDCYRGNKKNYMNLSQTSRGLTCSM		414
Db	342	LRMWYC-EIPSCSSASAPDQSDSVPRPEQTPVQEBEYQSDGQSRQTSSTITIGKKQCS		400
QY	415	WKNMEDLHRHIWE--PDASKINENYCNRPDDAHGPMCYGPNLIPWDYCPISCEG		471
Db	401	WAAMFP--HRHSKTPENFPDAG--LENNYCNRPDGD--KGPWCYTTPDSVREWCNMLKRC-S		455
QY	472	DTPTIYNLHPVYS-----		486
Db	456	ETGGSVEL--PVSQEPSCPSDETDCAITGNGKDYGKTAVTAAITPCGMAAOPERRH		513
QY	487	-----CAKTKOL-----		493
Db	514	SIFTPQNPRADELEKNYCNRPDQGVNCPMCTYTNPRKLYDCDIPICASASSPFCGRQY		573
QY	494	-----RVVNG-IPRTYINIGMNVSLRY--NKHIGGSLIKESWVLTAROCFSPRLKD		543
Db	574	EPKKCPGRVYGGCAVANHSPMWOJLSITRTFGQHFCCGTLIAPEWVLTAAHCLEKSSRPE		633
QY	544	-YEAMLGIHUVHGVDGDKCOVLNVSOLVYGPBESDVLKMLKAPAYALDQFVSTIDLPNY		602
Db	634	FYKVIILAHBEYIRGLD--VOEISVAKLILEPNRNDLALKLSRPATITDKVIPACLPSP		691
QY	603	GCTIPEKTSYVYGMYGTGLINTDGLRVANLHYIMNGEKSQSHIRGKVTULNESEICAGAE		662
Db	692	NYMVAADRTICYITGMGEBTQGFAGGRLEKQADLPYIEKNVCNRYEYLNVRKXSTELAGQL		751
QY	663	KIGSGPEBQYGGRLVDEQHKMRKAVLGYVPGRCALIPNRPGLFVRYVAYIAKYAKYHK		718
Db	752	AGGVDSQGGSGGGLVCFERDKYILQGVTSWGLGACAPRNKRGVVYVRSRVDWYIER		807

RESULT	9		
PLMN_HUMAN			
ID	PLMN_HUMAN	STANDARD;	PRT; 810 AA
AC	P00747;		
DT	21-JUL-1986	(rel. 01, Created)	
DT	01-MAR-1989	(rel. 10, Last sequence update)	

DE 16-OCT-2001 [Ref.: 40, Last annotation update)
PLasmimogen precursor (EC 3.4.21.7) [contains: Angiotatin].
CN Plg.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=90202879; PubMed=2318848;
RX Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
RL in the fibrinolytic system.";
RN J. Biol. Chem. 265:6104-6111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
RT for human plasminogen.";
RN FEBS Lett. 213:254-260(1987).
RN [3]
RP SEQUENCE OF 20-810.
RA Sottirup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the PIR data bank.
RN [4]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RL human and bovine plasminogen.";
RN Biochemistry 23:4243-4250(1984).
RN [5]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
RT of human plasminogen and their interaction with the NH2-terminal
Eur. J. Biochem. 50:489-494(1975).
RN [6]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottirup-Jensen L., Cleys H., Zajdel M., Petersen T.E., Magnusson S.;
RT (In) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
RL Progress in Chemical Fibrinolysis and Thrombolysis, pp.3:191-209,
RL Raven Press, New York (1978).
RN [7]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
RL plasminogen that forms the linkage between the plasmin chains.";
RN Eur. J. Biochem. 58:539-547(1975).
RN [8]
RP SEQUENCE OF 581-810.
RX MEDLINE=77225245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RN Eur. J. Biochem. 76:129-137(1977).
RN [9]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arradon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
RT of human plasmin: light (B) chain active center histidine sequence.";
RN J. Biol. Chem. 248:1631-1633(1973).
RN [10]
RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Groskopf W.R., Summaria L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
RT sequence of a peptide containing the active center serine residue.";
RN J. Biol. Chem. 244:3590-3597(1969).

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2002, 17:45:54 ; Search time 34.17 Seconds

(without alignments)
3685.698 Million cell updates/sec

Title: US-09-423-516-2

Perfect score: 4126
Sequence: 1 MMVTKLPALHQLVHLHL.....VAYYAKMKIKILTKVPS 728

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

SPTREMBL_19:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organella:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3880	94.0	728	6 Q9BH09	Q9BH09 felis silve
2	3170	76.8	726	13 Q90978	Q90978 gallus gall
3	2806	68.0	710	13 Q91402	Q91402 xenopus lae
4	1704	41.3	716	13 Q91691	Q91691 xenopus lae
5	1652.5	40.1	717	13 P70006	P70006 xenopus lae
6	1632	39.6	290	4 Q02935	Q02935 homo sapien
7	1627	39.4	704	13 Q90865	Q90865 gallus gall
8	1624	39.4	296	4 Q14519	Q14519 homo sapien
9	1615	39.1	716	11 P70521	P70521 rattus norv
10	1605	38.9	716	11 Q91XG8	Q91XG8 mus musculu
11	1579	38.3	709	13 Q90ZM6	Q90ZM6 brachydanto
12	1452	35.2	806	6 Q18783	Q18783 macropus eu
13	1401.5	34.0	812	11 Q9R0W3	Q9R0W3 rattus norv
14	1388.5	33.7	812	11 Q91WU5	Q91WU5 mus musculu
15	1371	33.2	810	4 Q15146	Q15146 homo sapien
16	1365	33.1	567	4 Q13208	Q13208 homo sapien

17	1364.5	33.1	648	4 Q9H1V4	Q9H1V4 homo sapien
18	1139	27.6	208	4 Q9BYM0	Q9BYM0 homo sapien
19	1139	27.6	210	4 Q13494	Q13494 homo sapien
20	1040	25.2	215	13 Q42341	Q42341 gallus gall
21	1036	25.1	211	11 Q55027	Q55027 mus musculu
22	945	22.9	2869	6 Q28398	Q28398 erinaceus e
23	785	19.0	454	6 Q46506	Q46506 papio hamad
24	654.5	15.9	313	13 Q9PU78	Q9PU78 crocodylus
25	615	14.9	109	6 Q9N1B8	Q9N1B8 ovis aries
26	556.5	13.5	334	6 Q46507	Q46507 papio hamad
27	555.5	13.5	286	4 Q13209	Q13209 homo sapien
28	549.5	13.3	559	11 Q91VP2	Q91VP2 mus musculu
29	523.5	12.7	516	4 Q98U99	Q98U99 homo sapien
30	500	12.1	607	13 Q91001	Q91001 gallus gall
31	494.5	12.0	608	13 Q9PTW7	Q9PTW7 struthio ca
32	471	11.4	597	11 Q35727	Q35727 mus musculu
33	462	11.2	616	6 Q97507	Q97507 sus scrofa
34	439	10.6	75	6 Q9BGN9	Q9BGN9 bos taurus
35	435.5	10.6	385	5 Q25101	Q25101 herdamia m
36	435	10.5	265	13 Q90WS3	Q90WS3 elaphe sp.
37	423.5	10.3	560	4 Q14520	Q14520 homo sapien
38	402	9.7	395	4 Q9BZM1	Q9BZM1 homo sapien
39	393	9.5	431	4 Q969W6	Q969W6 homo sapien
40	365	8.6	273	11 Q921N4	Q921N4 mus musculu
41	354	8.6	339	11 Q99L44	Q99L44 mus musculu
42	344.5	8.3	389	13 Q9PVX7	Q9PVX7 xenopus lae
43	340.5	8.3	643	6 Q97506	Q97506 sus scrofa
44	338.5	8.2	329	13 Q42272	Q42272 xenopus lae
45	338	8.2	275	4 Q96RZ6	Q96RZ6 homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	728 AA.
Q9BH09	Q9BH09			
AC	Q9BH09			
DT	01-JUN-2001 (T-REMBLrel. 17, Created)			
DT	01-DEC-2001 (T-REMBLrel. 19, Last sequence update)			
DE	HEPATOCTE GROWTH FACTOR HGF.			
OS	Felis silvestris catus (Cat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.			
OX	NCBI_TaxID=9685;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LIVER;			
RA	Kobayashi Y., Nakamura N., Ishizaka T., Masuda K., Ohno K.,			
RA	Tsujiimoto H.;			
RT	"Molecular cloning of feline hepatocyte growth factor (HGF) cDNA."			
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.			
DR	EMBL; AB046610; BAB21499.1; -			
DR	HSSP; P14210; 1BHT			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR000001; Kringle.			
DR	InterPro; IPR003014; PAN.			
DR	InterPro; IPR003609; Pan_app.			
DR	InterPro; IPR001254; Trypsin.			
DR	Pfam; PF00051; kringle; 1.			
DR	Pfam; PF00024; PAN; 1.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	PRINTS; PR00018; KRINGLE.			
DR	SMART; SM00130; KR; 4.			
DR	SMART; SM00473; PAN_AP; 1.			
DR	SMART; SM00020; TRYP_SPE; 1.			
DR	PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.			
DR	PROSITE; PS50070; KRINGLE_2; 4.			
DR	PROSITE; PS50240; TRYPSIN_DOM; 1.			

KW Hydrolase; Serine protease.
SQ SEQUENCE 728 AA: 83067 MW: 807F4A333D1E190A CRC64;

Query Match 94.0%; Score 3880; DB 6; Length 728;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 680; Conservative 28; Mismatches 18; Indels 4; Gaps 2;

```

QY 1 MMTVTLPLALLQVHLHLHLPIAIPYAEGRKRRNTIHERKSAKTTLIKIDPALKIK 60
DB 1 MMTVTLPLVLLQVHLHLHLHLPIAIPYAEGRKRRNTIHERKSAKTTLIKIDPALKIK 60
QY 61 TTKVNTADOCANRCKTRNKGGLPTCKAFVFDKARKQCLTFPNSMSGKKKEGHEFDLYE 120
DB 59 TTKMTADOCANRCKTRNKGGLPTCKAFVFDKARKQCLTFPNSMSGKKKEGHEFDLYE 118
QY 121 NKDYIRNCIIIGKRSYKGTASTITSGIKCOPWSSMTIHEHSFLPSSYRGKDLQENYCRNP 180
DB 119 NKDYIRNCIIIGKRSYKGTASTITSGIKCOPWSSMTIHEHSFLPSSYRGKDLQENYCRNP 178
QY 181 RGEEGPWCFTSNPEVREYVCDIPQCESEVEMTCNGESYRGDLMDHESGKICQRMHDQTP 240
DB 179 RGEEGPWCFTSNPEVREYVCDIPQCESEVEMTCNGESYRGDLMDHESGKICQRMHDQTP 238
QY 241 HRHKEFLPERYPDKGFDNYCRNPQGPWCYTLDPTHTWEYCAIKTCADNTMNDTDPVL 300
DB 239 HRHKEFLPERYPDKGFDNYCRNPQGPWCYTLDPTHTWEYCAIKTCADNTMNDTDPVL 298
QY 301 ETTECTIOGEGRGYVNTIWNIGIPQQRMDSOYPHEDMTPEWFKCKDLRENYCRNPDS 360
DB 299 ETTECTIOGEGRGYVNTIWNIGIPQQRMDSOYPHEDMTPEWFKCKDLRENYCRNPDS 358
QY 361 ESPWCETTPNINRGYCSOIPNCDMSHGDDCYRGKSKNTMGLSOTRSLGTCMMDKNME 420
DB 359 ESPWCETTPNINRGYCSOIPNCDMSHGDDCYRGKSKNTMGLSOTRSLGTCMMDKNME 418
QY 421 DLHRHTEMERPDASKLNEYCRNPDDAHGPMCYTNPILPMDYCPISCEEDDTPPTVNL 480
DB 419 DLHRHTEMERPDASKLNEYCRNPDDAHGPMCYTNPILPMDYCPISCEEDDTPPTVNL 478
QY 481 DHPVISCATKQOLRVNIGIPRTNIGMMVSLRYRKATICGSLIKESVLTAROCFSPSR- 539
DB 479 DHPVISCATKQOLRVNIGIPRTNIGMMVSLRYRKATICGSLIKESVLTAROCFSPSR- 538
QY 540 -DLKRYEAMIGIHVHGDEKCKQOVNLVSQLVYGPESGSDLYLKLRLPAVLDDFVSTID 598
DB 539 KDLKRYEAMIGIHVHGDEKCKQOVNLVSQLVYGPESGSDLYLKLRLPAVLDDFVSTID 598
QY 599 LPNVCCTIPEKTSYVGMGYTGLINVDGLIRVAHLIYMGNEKCSQHHRGKVTLINESEIC 658
DB 599 LPNVCCTIPEKTSYVGMGYTGLINVDGLIRVAHLIYMGNEKCSQHHRGKVTLINESEIC 658
QY 659 AGAKTIGSGPEEGYGPPLVCEQHKRMVLTGYIVPGGCAIPNRPGLFVAVAYAKMIHK 718
DB 659 AGAKTIGSGPEEGYGPPLVCEQHKRMVLTGYIVPGGCAIPNRPGLFVAVAYAKMIHK 718
QY 719 IILTYKVPQS 728
DB 719 IILTYKVPQS 728
QY 719 IILTYKIPQS 728
DB 719 IILTYKIPQS 728

```

RESULT 2
090978 PRELIMINARY: PRT: 726 AA.
AC 090978: 090866;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE HEPATOCYTE GROWTH FACTOR / SCATTER FACTOR PRECURSOR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.

```

OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RX MEDLINE=96029010; PubMed=7554499;
RA Thery C., Sharpe M.D., Bailey S.D., Stern C.D., Gherardi E.;
RT "Expression of HGF/SF, HGF/MSP, and c-met suggests new functions
RL during early chick development.";
RN [2]
RP SEQUENCE OF 1-409 FROM N.A.
RC STRAIN=WHITE LEGHORN, AND RHODE ISLAND RED X LIGHT SUSSEX;
RC TISSUE=EMBRYO;
RX MEDLINE=95237013; PubMed=7720585;
RA Streit A., Stern C.D., Thery C., Ireland G.W., Aparicio S.,
RT Sharpe M.J., Gherardi E.;
RT "A role for HGF/SF in neural induction and its expression in Hensen's
RL node during gastrulation.";
CC Development 121:813-824(1995).
-i SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY
DR EMBL: X84045; CAAS8864.1; -
DR HSP: X80131; CAAS6430.1; -
DR HSP: P14210; IBHT.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR Pfam: IPR001254; Trypsin.
DR Pfam: PR00051; Kringle; 4.
DR Pfam: PR00024; PAN; 1.
DR Pfam: PR00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_AP. 1.
DR SMART: SM00020; Tryp_Spec; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS0070; KRINGLE_2; 4.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
KW Alternative splicing; Glycoprotein; Growth factor; Hydrolase; Kringle;
KW Serine protease; Serine protease homolog; Signal.
FT SIGNAL 1 27
FT CHAIN 28 726 HEPATOCYTE GROWTH FACTOR/ SCATTER FACTOR.
SQ SEQUENCE 726 AA: 82913 MW: 5805F048A576C38 CRC64;

```

Query Match 76.8%; Score 3170; DB 13; Length 726;
Best Local Similarity 74.5%; Pred. No. 1,4e-280;
Matches 544; Conservative 92; Mismatches 86; Indels 8; Gaps 3;

```

QY 1 MMTVTLPLALLQVHLHLHLPIAIPYAEGRKRRNTIHERKSAKTTLIKIDPALKIK 60
DB 1 MMTVTLPLVLLQVHLHLHLHLPIAIPYAEGRKRRNTIHERKSAKTTLIKIDPALKIK 60
QY 61 TTKVNTADOCANRCKTRNKGGLPTCKAFVFDKARKQCLTFPNSMSGKKKEGHEFDLYE 120
DB 57 TTKVNTADOCANRCKTRNKGGLPTCKAFVFDKARKQCLTFPNSMSGKKKEGHEFDLYE 116
QY 121 NKDYIRNCIIIGKRSYKGTASTITSGIKCOPWSSMTIHEHSFLPSSYRGKDLQENYCRNP 180
DB 117 NKDYIRNCIIIGKRSYKGTASTITSGIKCOPWSSMTIHEHSFLPSSYRGKDLQENYCRNP 176
QY 181 RGEEGPWCFTSNPEVREYVCDIPQCESEVEMTCNGESYRGDLMDHESGKICQRMHDQTP 240
DB 177 RGEEGPWCFTSNPEVREYVCDIPQCESEVEMTCNGESYRGDLMDHESGKICQRMHDQTP 236
QY 241 HRHKEFLPERYPDKGFDNYCRNPQGPWCYTLDPTHTWEYCAIKTCADNTMNDTDPVL 300
DB 237 HRHKEFLPERYPDKGFDNYCRNPQGPWCYTLDPTHTWEYCAIKTCADNTMNDTDPVL 296
QY 301 ETTECTIOGEGRGYVNTIWNIGIPQQRMDSOYPHEDMTPEWFKCKDLRENYCRNPDS 360

```

```
Db 297 ETTTCIOGGEYRGVNTWISGICQRMDSOPPHNITPENFKCKDLRENYCRNDGS 356
QY 361 ESPWCFTTDNINRGYCSQIPNCDMSHGQDCYRGNGKNYGNLSQTSGLTCSMDKNME 420
Db 357 ESPWCFTTDNINRGYCSQIPNCDVSNEDOCYRGNGKSYGMNLSKTRFGLTCSMDKNIE 416
QY 421 DLHRH--IFMEPDASKINENYCRNPDDAHGPMCYTGNPLIPMPCYISRCEDGTTPTIY 478
Db 417 DLRRHIOIFRBPVSKKKNYCRNPDDHFGPMCYTDDPLIPMPCYISRCGTTPTTTT 476
QY 479 NLDPVISCARKTQLRVNGVNPRTNIGMVSILRYRNKHICGSLIKESWVLAAROCFPS 538
Db 477 SLDDTVPICASTKHLRVNVPITQTEGKGVVSLYRNKHICGSLIKESWVLAAROCFPS 536
QY 539 R--DLKRYEAMLGIDHVGHGDECKQVLAIVSOLVYVPEBESDVLMLKLARAVLDDEVST 596
Db 537 RYDLDKRYKAMLGVHNKKGEGEKNROVNRINSOLVYVPAQSDVLVLLSRPAITLTFVEI 596
QY 597 IDLPNYGCTIPERTSCSVYGMVGTGLINVDGLLRVAHLVIMNGEKCSQHHRGKVTLINESE 656
Db 597 IRLPISGCTIPERTSCSVFGMGTGLRPNYDGLLRVANLFLIGNKCKNOYLKTTVNESE 656
QY 657 ICAGAETIGAGPCERDYGGLVCEQNRKLIYGVIVPGRCALIRNPGRGIFVRYAVAKWI 716
Db 657 ICAVAETIGAGPCERDYGGLVCEQNRKLIYGVIVPGRCALIRNPGRGIFVRYAVAKWI 716
QY 717 HKIITPYKVP 726
Db 717 HKIMTYRKRP 726

RESULT 3
091402 PRELIMINARY; PRT; 710 AA.
AC 091402;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE HEPATOCYTE GROWTH FACTOR.
GN HGF.
OS Xenopus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodinae.
OX NCBI_TaxID=8353;
RN [1]
SEQUENCE FROM N.A.
TISSUE-TAILBUD;
MEDLINE=95267690; PubMed=7748783;
Nakamura H., Tashiro K., Nakamura T., Shiohawa K.;
"molecular cloning of Xenopus HGF cDNA and its expression studies in
Xenopus early embryogenesis.";
Mech. Dev. 49:123-131(1995).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY.
EMBL: S77422; AAB34354.2; -.
DR HSSP: P14210; 1BHT.
DR MEROPS: S01.976; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Trypsin.
DR Pfam: PF00051; Kringle; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_AP; 1.
DR SMART: SM00020; TRYSPC; 1.
DR PROSITE: PS00021; KRINGLE_1; 3.
DR PROSITE: PS50070; KRINGLE_2; 4.
```

```
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 710 AA; 81487 MW; 5FE640BE31C27FC CRC64;

Query Match 68.0%; Score 2806; DB 13; Length 710;
Best Local Similarity 66.1%; Pred. No. 2,4e-247;
Matches 481; Conservative 107; Mismatches 120; Indels 20; Gaps 5;

QY 1 MWTYKLLPALQLHVLHLLPLAIPYAEQGRKRRNTIEFFKSAKTLIKIDPALKIK 60
Db 1 MWTYKCAVIFEL-----ITLAEGRGKRRAAFDDYKKAETTLRLKALEVK 47
QY 61 TKRYNTDOCANRCTRGKLPFTCKAFVPEDKAROCIMFPENSSGVCARKEGHEPYLE 120
Db 48 TKMFNTENCAKRSRNGKLPFTCKAFEDBNIRCHMFENISAGIKDYDISFDLYE 107
QY 121 NKDYIRNCIIGKRSYGYTSITFSGIKQCPWSSMIPHEHSFLPSYRGKDLQENYCRNP 180
Db 108 KKDYIRNCIIGKRSYGYTSITFSGIKQCPWSSMIPHEHSFLPSYRGKDLQENYCRNP 167
QY 181 RGEFGWPCFTSNPEVAREVCDIPQCEVBCMTNGESYRGMDHTSGKICQMDHOTP 240
Db 168 KGEFGWPCFTSNPEVAREVCDIPQCEVBCMTNGESYRGMDHTSGKICQMDHOTP 227
QY 241 HRHKFLPERYPDKGFDDNYCRNPDPGPRMPCYTLDPHTRMRYCAIKTCADNTMNDTVP 300
Db 228 HRHKFLPERYPDKGFDDNYCRNPDPGPRMPCYTLDPHTRMRYCAIKTCADNTMNDTVP 285
QY 301 ETTTCIOGGEYRGVNTWISGICQRMDSQYVPEHDMPEPFENFKCKDLRENYCRNDGS 360
Db 286 -TKDCMGQGGYRGVSVSTYNGIQCQRMDSQFPHLHNFPEFKCKDLRENYCRNDGS 344
QY 361 ESPWCFTTDNINRGYCSQIPNCDMSHGQDCYRGNGKNYGNLSQTSGLTCSMDKNME 420
Db 345 ESPWCFTTDNINRGYCSQIPNCDMSHGQDCYRGNGKNYGNLSQTSGLTCSMDKNME 403
QY 421 DLHRHIFMEPDASKINENYCRNPDDAHGPMCYTGNPLIPMPCYISRCEDGTTPTIYN 480
Db 405 DLKRYEAMLGIDHVGHGDECKQVLAIVSOLVYVPEBESDVLMLKLARAVLDDEVST 464
QY 481 DHPVISCARKTQLRVNGVNPRTNIGMVSILRYRNKHICGSLIKESWVLAAROCFPSR 539
Db 465 DHPVISCARKTQLRVNGVNPRTNIGMVSILRYRNKHICGSLIKESWVLAAROCFPSR 523
QY 540 -DLKRYEAMLGIDHVGHGDECKQVLAIVSOLVYVPEBESDVLMLKLARAVLDDEVST 598
Db 524 IDLYEAMLGIDHVGHGDECKQVLAIVSOLVYVPEBESDVLMLKLARAVLDDEVST 582
QY 599 LPNYGCTIPERTSCSVYGMVGTGLINVDGLLRVAHLVIMNGEKCSQHHRGKVTLINESE 658
Db 583 LPNYGCTIPERTSCSVYGMVGTGLINVDGLLRVAHLVIMNGEKCSQHHRGKVTLINESE 642
QY 659 AGAETIGAGPCERDYGGLVCEQNRKLIYGVIVPGRCALIRNPGRGIFVRYAVAKWI 718
Db 643 AIGETANIGPCERDYGGLVCEQNRKLIYGVIVPGRCALIRNPGRGIFVRYAVAKWI 702
QY 719 IITPYKVP 726
Db 703 IIMTYKAP 710

RESULT 4
091691 PRELIMINARY; PRT; 716 AA.
AC 091691;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE GROWTH FACTOR LIVERLINE.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
```


OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ruiz i Altaba A., Thery C.,
 RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U57455; AAB52574.1; -
 DR HSSP: P00747; ICEA.
 DR MEROPS: S01.00A; -
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR Pfam: PF00051; Kringle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00018; KRINGLE.
 DR SMART: SM00130; KR; 4.
 DR SMART: SM00473; PAN_AP; 1.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00021; KRINGLE_1; 4.
 DR PROSITE: PS00070; KRINGLE_2; 4.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 KM Hydrolyase; Serine protease
 SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;

Query Match 41.3%; Score 1704; DB 13; Length 716;
 Best Local Similarity 44.5%; Pred. No. 9, 8e-147;
 Matches 310; Conservative 124; Mismatches 236; Indels 24; Gaps 12;

OY 36 RNTIHEFKSAATTLKIDIPALKIKKKVNTDQCANRCKTNNKGLPTCKAFVDFKARKQ 95
 Db 31 RSLALNDYQSKGLELVHNNG-GVKQEIQSEIYQVCAKQSD---LIDRSFVYMMKSGT 85
 OY 96 CLMFEPNSMSSGVKKEFEHFDLYENKDYIRNCIIIGKGSYKGVSTIRKSGIKCQPMISM 155
 Db 86 CRLLPWTQNSAVNLQRRNQYDLQKQDIYRCQVANGNTVIGTSKTSKSGTCQRMRLK 145
 OY 156 ITHHSFLPSSIRKGLDQENYCRNPGEGGPMCTSNPEVRYEVCIDIPOGSEVECMTCN 215
 Db 146 FPHDHKFSPIHW--PELEENYCRNPDSPDEGPMCTTQKNIHQVCGIKKCEDAVCLCN 203
 OY 216 GESYGLMDHTESGKICQICQMDHQTQTHRKFLPERYPDGFQDNVCRNPDGQPRPCYILD 275
 Db 204 GEDYGSVDRTSSGKECCQRMQDQTHAPRYKPKTPDPSLDNVCNRPDSSERPMCTYTD 263
 OY 276 PHTREYCAIKTCADNTMNDTVPLETTTCIQQGEGYRGTVNTIWNIGIPCQRMDSQYPH 335
 Db 264 PVEKEFCRIKCKKQRLSNIEI--TSTCFKREGEYRGKANTTTSIGIPQRMDSQTPQ 320
 OY 336 EHDMPENFKCKDLRENYCRNPDGSESPWCTTDPNINVGICSOIPNC-DMSHQDCYRG 394
 Db 321 SHRLPEYKPKGLDENYCRNPDGSEAPWCFTTLPGRMAVCFQIKRCKDQVLEPDCYHG 380
 OY 395 NGKNVYMCNLSQTRSLTCSMDKMMEDLRAHIFWEPDASKNENYCRNPDGADGPMPCYT 454
 Db 381 NGELYSGVSKTRGKICRMRWEKRNDELSD-QPLVLPLEENYCRNPDGSDGPMPCYT 439
 OY 455 GNPRLPMDYCIISRCEDTPTIYNLDHPIV-SCAKT----KQLRVNNGIPTPTNIGMM 508
 Db 440 MDPMPDPDYCAIKPCBEKEVLLTEEAESIYFDSGCRNDRIFQKSRILVGPV--GNSPWT 497
 OY 509 VSLNVR-NKNIICGSSLIKESVNLAROCFPS--RDLDYEMALGIHNVHGRDEKCKOYL 565
 Db 498 VSLNRNGEHNCGSLVKKENNVISTRCFSSCDADLSGYEAVMQLTKPNSPDPDRKQSV 557
 OY 566 NVSQLVYGPESGDLVLMKLARPAVLDFVSTIDLPNPGCTIPKTSQSVYGVGTGLIN 625
 Db 558 PINKIVGSPDSSLVMLKLERPITLNSRVALICLPPRIYVPEITCKEIGAKWDTRGTGH 617
 OY 626 DGLRLVNLKIMGNEKSOHNRG-KVTLNSETICAGAEKIGSGPCEGDYGGRLVCEQHKM 684

Db 618 DNLKIAIEFYIISDNCNKNYRSQQRNKVSDNEMCTKPMPIDAGAGEGYGGLACTHDC 677
 OY 685 RAVLGIVPGRCALPNRPGIYFRAVYAKRTHKII 720
 Db 678 LVLEGIVPARCGCKKNQPAIFRVSIVYDWINKV 713
 RESULT 5
 ID P70006 PRELIMINARY; PRT; 717 AA.
 AC P70006;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN PRECURSOR.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesodactylia; Pipridae; Pipridae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=96404125; PubMed=8808403;
 RA Aberger F., Schmidt G., Richter K.;
 RT "The Xenopus homologue of hepatocyte growth factor-like protein is specifically expressed in the presumptive neural plate during gastrulation."
 RL Mech. Dev. 54:23-37(1996).
 DR EMBL: Y08734; CAA69989.1; -
 DR HSSP: P00747; ICEA.
 DR MEROPS: S01.977; -
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR Pfam: PF00051; Kringle; 4.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00018; KRINGLE.
 DR SMART: SM00130; KR; 4.
 DR SMART: SM00473; PAN_AP; 1.
 DR SMART: SM00020; Tryp_Spc; 1.
 DR PROSITE: PS00021; KRINGLE_1; 3.
 DR PROSITE: PS50070; KRINGLE_2; 4.
 DR PROSITE: PS50240; TRYPSIN_DOM; 1.
 DR Hydrolyase; Serine protease; Signal.
 FT SIGNAL 1 28
 FT CHAIN 29 717 HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.
 FT SEQUENCE 717 AA; 82017 MW; 6F877A432C8CD54 CRC64;

Query Match 40.1%; Score 1652.5; DB 13; Length 717;
 Best Local Similarity 42.8%; Pred. No. 4, 9e-142;
 Matches 306; Conservative 126; Mismatches 258; Indels 25; Gaps 13;

OY 17 LHLPLPIAIPYAEQGRKRRNTIHEFKSAATTLKIDIPALKIKKKVNTAQCANRCTR 76
 Db 14 LFLCFNAVAFVLTG-TAHRSLANDYQSKGLELVHNNG-GVKQEVQAEIYQICAKQSD 71
 OY 77 NKLPLPTCAFYFDKARKKQCLLPENFMSGVKKEGHEFDLYENKDYIRNCIIIGKSGY 136
 Db 72 ----LIDCRSFDYMSQSCRLLPMTQNSVNLQRRNQYDLQKQDIYRCQVANGNTY 127
 OY 137 KGVSTIKSGICQPMSSNIPHEHSLPSSYSGKLDQENYCRNPGEGGPMCTSNPEV 196
 Db 128 RGVSTKTKNGRCIQHRLKLPDHKFSPTHW--PELEENYCRNPDSPDEGLMCTYTDKNI 185
 OY 197 RYEVCDIPQSEVECMTCNNGESYGLMDHTESGKICQICQMDHQTQTHRKFLPERYPDGF 256
 Db 186 RHQYCGIKKCEDAVCLTNGEDYRGSVDRTSSGKECCQRMQDQAPHTHRYKPKTPDPSLD 245
 OY 257 DNYCRNPDGQPRPCYITDTPHTREYCAIKTCADNTMNDTVPLETTTCIQQGEGYRG 316

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2002, 17:43:19 ; Search time 14.56 Seconds
(without alignments)
1221.280 Million cell updates/sec

Title: US-09-423-516-2

Perfect score: 4126

Sequence: 1 MMTVRLPALLLQHVLLHLL.....VAVYAKWIKIKITLVKVPQS 728

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Indexed: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PTCUS.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfilest.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4126	100.0	728	1	US-07-815-333A-2
2	4126	100.0	728	1	US-08-087-783A-22
3	4088.5	99.1	723	1	US-07-838-410-1
4	4088.5	99.1	723	1	US-08-290-937B-1
5	4088.5	99.1	723	1	US-08-404-643-1
6	4088.5	99.1	723	1	US-09-194-346-1
7	4085.5	99.0	723	1	US-08-290-937B-2
8	4079.5	98.9	723	1	US-08-290-937B-3
9	4077.5	98.8	723	4	US-08-030-410-3
10	4064.5	98.5	723	4	US-09-194-376-2
11	4064.5	98.5	723	4	US-09-194-326-3
12	3968	96.2	697	2	US-08-460-890A-50
13	3968	96.2	697	2	US-08-167-641C-50
14	3968	96.2	697	4	US-08-460-971A-50
15	3968	96.2	697	4	US-08-463-040-50
16	1654	40.1	711	2	US-08-766-982-2
17	1654	40.1	711	2	US-09-296-219-2
18	1652	40.0	711	1	US-08-184-012C-8
19	1652	40.0	711	1	US-08-334-177-2
20	1652	40.0	711	2	US-08-666-082B-1
21	1652	40.0	711	5	PCT-US95-13830-2
22	1608	39.0	716	2	US-08-766-982-1
23	1608	39.0	716	4	US-09-296-219-1
24	1379.5	33.4	812	1	US-08-248-629A-1
25	1379.5	33.4	812	1	US-08-451-932-1
26	1379.5	33.4	812	1	US-08-453-260-1
27	1379.5	33.4	812	1	US-08-326-785-1

28	1379.5	33.4	812	2	US-08-612-788-1	Sequence 1, Appli
29	1379.5	33.4	812	2	US-08-605-598B-1	Sequence 1, Appli
30	1379.5	33.4	812	2	US-08-429-743-1	Sequence 1, Appli
31	1379.5	33.4	812	2	US-08-866-735-1	Sequence 1, Appli
32	1379.5	33.4	812	3	US-09-066-028-1	Sequence 1, Appli
33	1379.5	33.4	812	5	PCT-US95-05107-1	Sequence 1, Appli
34	1372.5	33.3	791	3	US-08-643-219-1	Sequence 1, Appli
35	1372.5	33.3	791	3	US-08-851-350-1	Sequence 1, Appli
36	1372	33.3	790	1	US-08-469-486-54	Sequence 54, Appli
37	1372	33.3	790	2	US-08-469-658-54	Sequence 54, Appli
38	1370	33.2	810	1	US-07-854-603-2	Sequence 29, Appli
39	1370	33.2	810	4	US-08-147-000B-29	Sequence 29, Appli
40	1370	33.2	814	1	US-09-086-514-1	Sequence 1, Appli
41	1370	33.2	814	1	US-08-750-711-1	Sequence 1, Appli
42	1367.5	33.1	791	2	US-09-131-995-1	Sequence 1, Appli
43	1367.5	33.1	791	2	US-08-832-087B-1	Sequence 1, Appli
44	1367.5	33.1	791	4	US-09-132-154-1	Sequence 1, Appli
45	1366	33.1	810	6	5200340-8	Patent No. 5200340

ALIGNMENTS

RESULT 1
US-07-815-333A-2

Sequence 2, Application US/07815333A

Patent No. 5342831

GENERAL INFORMATION:

APPLICANT: Nakamura, Toshikazu

APPLICANT: Matsumoto, Kunio

TITLE OF INVENTION: EPITHELIOCYTE GROWTH ACCELERATOR

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Leydig, Volt & Mayer

STREET: Two Prudential Plaza, Suite 4900

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/815,333A

FILING DATE: 19911227

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Green, Robert F.

REGISTRATION NUMBER: 27555

REFERENCE/DOCKET NUMBER: 44069

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 616-5600

TELEFAX: (312) 616-5700

TELEX: 25-3533

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 728 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-815-333A-2

Query Match 100.0%; Score 4126; DB 1; Length 728;

Best Local Similarity 100.0%; Pred. No 0;

Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMTVRLPALLLQHVLLHLLPIAIPYAEGRKRRNTIHEKKSATKTLIKIDPALKIK 60
Db 1 MMTVRLPALLLQHVLLHLLPIAIPYAEGRKRRNTIHEKKSATKTLIKIDPALKIK 60

```

QY 61 TRKVNADOCANCRTRNGLPFTCKAFVFDKARKOCLMEFPNMSGGVKKKEFEFDLYE 120
DB 61 TRKVNADOCANCRTRNGLPFTCKAFVFDKARKOCLMEFPNMSGGVKKKEFEFDLYE 120
QY 121 NKDYINCLIGGRSYKGTAVSTKSGIKQCPMSSMIPHEHSFLPSSYRGKDLQENYCRNP 180
DB 121 NKDYINCLIGGRSYKGTAVSTKSGIKQCPMSSMIPHEHSFLPSSYRGKDLQENYCRNP 180
QY 181 RGEEGPWCFTSNPEVRYEVCIDPQSEVECMTCNGESTRGMLDHTESKICQRMWQHP 240
DB 181 RGEEGPWCFTSNPEVRYEVCIDPQSEVECMTCNGESTRGMLDHTESKICQRMWQHP 240
QY 241 HRHKLPERYPDKGEGDNYCRNPDGQPRPWCYTLDPHTMEYCAIKTCADNTMNDTVL 300
DB 241 HRHKLPERYPDKGEGDNYCRNPDGQPRPWCYTLDPHTMEYCAIKTCADNTMNDTVL 300
QY 301 ETTECIQGGEGYRGVNTINMGICQQRWDSQYRPHEDMTPEHFCKDLRENYCRNPDGS 360
DB 301 ETTECIQGGEGYRGVNTINMGICQQRWDSQYRPHEDMTPEHFCKDLRENYCRNPDGS 360
QY 361 ESPMCFETDPNIRVGYCSQIPNCDMSHGDDCYRGNGKNYMGNLISOTRSGLTCSMDKNME 420
DB 361 ESPMCFETDPNIRVGYCSQIPNCDMSHGDDCYRGNGKNYMGNLISOTRSGLTCSMDKNME 420
QY 421 DLHRIHFWEPDASKLNEYCNCRNPDGDAHGPWCYTGPNLPMDYCPISRCGDDTPTIVNL 480
DB 421 DLHRIHFWEPDASKLNEYCNCRNPDGDAHGPWCYTGPNLPMDYCPISRCGDDTPTIVNL 480
QY 481 DHPVISCACKTQOLRVVNGIPRTNIGMWVSLRYRNKHICGSLIKESVWLTARQCFPSRD 540
DB 481 DHPVISCACKTQOLRVVNGIPRTNIGMWVSLRYRNKHICGSLIKESVWLTARQCFPSRD 540
QY 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDYLMLKARPAVLDDFVSTIDLP 600
DB 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDYLMLKARPAVLDDFVSTIDLP 600
QY 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYMGNEKCSQHHRKQVTLNESEICAG 660
DB 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYMGNEKCSQHHRKQVTLNESEICAG 660
QY 661 AEKISGPGCEBDYGGLVCEQHKKRMVGLVIVPGRGCAIPNRPGLIFVAVYAKWIIHII 720
DB 661 AEKISGPGCEBDYGGLVCEQHKKRMVGLVIVPGRGCAIPNRPGLIFVAVYAKWIIHII 720
QY 721 LTYKVPQS 728
DB 721 LTYKVPQS 728

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/884811
FILING DATE: 18-MAY-92
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/885971
FILING DATE: 18-MAY-92
ATTORNEY/AGENT INFORMATION:
NAME: Marschang, Diane L.
REGISTRATION NUMBER: 35,600
REFERENCE/DOCKET NUMBER: P0755779P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-5416
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 728 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-087-783A-22

Query Match      100.0%; Score 4126; DB 1; Length 728;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMYTLRALLQHVLLHLLPLPIPYAEGQRKRRTIHEFKSAKTTLIKIDPALKIK 60
DB 1 MMYTLRALLQHVLLHLLPLPIPYAEGQRKRRTIHEFKSAKTTLIKIDPALKIK 60
QY 61 TRKVNADOCANCRTRNGLPFTCKAFVFDKARKOCLMEFPNMSGGVKKKEFEFDLYE 120
DB 61 TRKVNADOCANCRTRNGLPFTCKAFVFDKARKOCLMEFPNMSGGVKKKEFEFDLYE 120
QY 121 NKDYINCLIGGRSYKGTAVSTKSGIKQCPMSSMIPHEHSFLPSSYRGKDLQENYCRNP 180
DB 121 NKDYINCLIGGRSYKGTAVSTKSGIKQCPMSSMIPHEHSFLPSSYRGKDLQENYCRNP 180
QY 181 RGEEGPWCFTSNPEVRYEVCIDPQSEVECMTCNGESTRGMLDHTESKICQRMWQHP 240
DB 181 RGEEGPWCFTSNPEVRYEVCIDPQSEVECMTCNGESTRGMLDHTESKICQRMWQHP 240
QY 241 HRHKLPERYPDKGEGDNYCRNPDGQPRPWCYTLDPHTMEYCAIKTCADNTMNDTVL 300
DB 241 HRHKLPERYPDKGEGDNYCRNPDGQPRPWCYTLDPHTMEYCAIKTCADNTMNDTVL 300
QY 301 ETTECIQGGEGYRGVNTINMGICQQRWDSQYRPHEDMTPEHFCKDLRENYCRNPDGS 360
DB 301 ETTECIQGGEGYRGVNTINMGICQQRWDSQYRPHEDMTPEHFCKDLRENYCRNPDGS 360
QY 361 ESPMCFETDPNIRVGYCSQIPNCDMSHGDDCYRGNGKNYMGNLISOTRSGLTCSMDKNME 420
DB 361 ESPMCFETDPNIRVGYCSQIPNCDMSHGDDCYRGNGKNYMGNLISOTRSGLTCSMDKNME 420
QY 421 DLHRIHFWEPDASKLNEYCNCRNPDGDAHGPWCYTGPNLPMDYCPISRCGDDTPTIVNL 480
DB 421 DLHRIHFWEPDASKLNEYCNCRNPDGDAHGPWCYTGPNLPMDYCPISRCGDDTPTIVNL 480
QY 481 DHPVISCACKTQOLRVVNGIPRTNIGMWVSLRYRNKHICGSLIKESVWLTARQCFPSRD 540
DB 481 DHPVISCACKTQOLRVVNGIPRTNIGMWVSLRYRNKHICGSLIKESVWLTARQCFPSRD 540
QY 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDYLMLKARPAVLDDFVSTIDLP 600
DB 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYVGPESDYLMLKARPAVLDDFVSTIDLP 600
QY 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYMGNEKCSQHHRKQVTLNESEICAG 660
DB 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYMGNEKCSQHHRKQVTLNESEICAG 660
QY 661 AEKISGPGCEBDYGGLVCEQHKKRMVGLVIVPGRGCAIPNRPGLIFVAVYAKWIIHII 720
DB 661 AEKISGPGCEBDYGGLVCEQHKKRMVGLVIVPGRGCAIPNRPGLIFVAVYAKWIIHII 720

```

OY 721 LTYKVPOS 728
|||||
Db 721 LTYKVPOS 728

RESULT 3

US-07-838-410-1
; Sequence 1, Application US/07838410
; Patent No. 5328836
; GENERAL INFORMATION:
; APPLICANT: SHIMA, NOBUYUKI
; APPLICANT: HIGASHIO, KANJI
; APPLICANT: NAGAO, MASAYA
; APPLICANT: OOGAKI, FUMIKO
; APPLICANT: TAKAKA, HIROAKI
; APPLICANT: TSUDA, EISUKE
; TITLE OF INVENTION: PLASMIDS CONTAINING DNA ENCODING THE
; TITLE OF INVENTION: AMINO ACID SEQUENCE OF TCF-II, TRANSFORMED CELLS WITH THE
; TITLE OF INVENTION: PLASMIDS, AND PRODUCTION METHOD OF BIOLOGICALLY ACTIVE
; TITLE OF INVENTION: SUBSTANCE USING THE TRANSFORMED CELLS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD.
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/838,410
; FILING DATE: 19920311
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/00942
; FILING DATE: 15-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 185852-1990
; FILING DATE: 13-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: WHITE, JOHN L.
; REGISTRATION NUMBER: 17,746
; REFERENCE/DOCKET NUMBER: WAK 110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; TELEX: 64191
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL TYPE: Fibroblast
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 393..405
; OTHER INFORMATION: /note= "INTERNAL AMINO ACID
; OTHER INFORMATION: SEQUENCE IN ALPHA-CHAIN"
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 490..505
; OTHER INFORMATION: /note= "N-TERMINAL AMINO ACID
; OTHER INFORMATION: SEQUENCE OF BETA-CHAIN"
; FEATURE:

; NAME/KEY: Domain
; LOCATION: 605..623
; OTHER INFORMATION: /note= "INTERNAL AMINO ACID
; OTHER INFORMATION: SEQUENCE IN BETA-CHAIN"
US-07-838-410-1

Query Match 99.1%; Score 4088.5; DB 1; Length 723;
Best local Similarity 99.3%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

OY 1 MWYTKLLPALLQHVLLHLLPLAIPYAGQRRRTTIEFKSAKTTLIKIDPALIK 60
Db 1 MWYTKLLPALLQHVLLHLLPLAIPYAGQRRRTTIEFKSAKTTLIKIDPALIK 60
OY 61 TKRYNTADOCANRCTRRKGLPFTCKAFVFPKARKQCLMFPNSSGVKKFEGHEFDLYE 120
Db 61 TKRYNTADOCANRCTRRKGLPFTCKAFVFPKARKQCLMFPNSSGVKKFEGHEFDLYE 120
OY 121 NKDYIRKCIIGKSGSYGTVSITKSGIKQPMSSMIPHEHSFLPSSYRGKDLOENYCRNP 180
Db 121 NKDYIRKCIIGKSGSYGTVSITKSGIKQPMSSMIPHEH-----SYRGKDLOENYCRNP 175
OY 181 RGEEGPWCFTSNPEVRYEVCDDIPQCEVECMTCNGESYRGKMDHTESGKICQRMHDQTP 240
Db 176 RGEEGPWCFTSNPEVRYEVCDDIPQCEVECMTCNGESYRGKMDHTESGKICQRMHDQTP 235
OY 241 HRHKFLPERYPDKGFDNRYCRNPDPGQRPWCYTITDPTTRMEYCAIKTCADNTMNDTVPL 300
Db 236 HRHKFLPERYPDKGFDNRYCRNPDPGQRPWCYTITDPTTRMEYCAIKTCADNTMNDTVPL 295
OY 301 ETTFCIOGSGEGYGTNTVITWNGIPCOQRMDSQYRHEHDMTPENFKCKDLRENYCRNPDGS 360
Db 296 ETTFCIOGSGEGYGTNTVITWNGIPCOQRMDSQYRHEHDMTPENFKCKDLRENYCRNPDGS 355
OY 361 ESPWCFETDPIRIVGYCSQIIPNCDSHGDQCYRGNKRNMGNTLSQTRSGLTCSMMDKME 420
Db 356 ESPWCFETDPIRIVGYCSQIIPNCDSHGDQCYRGNKRNMGNTLSQTRSGLTCSMMDKME 415
OY 421 DLNRHIFWEPDASKLNNYCRNPDDDAHGPMCTYGNPLIPMDYCPISRCBDDTPTTVNL 480
Db 416 DLNRHIFWEPDASKLNNYCRNPDDDAHGPMCTYGNPLIPMDYCPISRCBDDTPTTVNL 475
OY 481 DHPVISCATKQRLRVVNGIIPRTNIGMWSLRYRNKHICGSSLIKESWVLTAROCFSPSRD 540
Db 476 DHPVISCATKQRLRVVNGIIPRTNIGMWSLRYRNKHICGSSLIKESWVLTAROCFSPSRD 535
OY 541 LKDYEAMLGIDVHGRDDEKQVLANVSOLVYGPESGDLVLMKLARPAVLDDFVSTIDLP 600
Db 536 LKDYEAMLGIDVHGRDDEKQVLANVSOLVYGPESGDLVLMKLARPAVLDDFVSTIDLP 595
OY 601 NYGCTIPEKTSQSYGNGYGLINVDGLLRVAHLIYINGNEKCSQHHKGYTLNSETICAG 660
Db 596 NYGCTIPEKTSQSYGNGYGLINVDGLLRVAHLIYINGNEKCSQHHKGYTLNSETICAG 655
OY 661 AEKIGSGPCBSGYGGLVCEQHKRMVLYIVPBGCAIPRPGIFVAVYAAWIKHII 720
Db 656 AEKIGSGPCBSGYGGLVCEQHKRMVLYIVPBGCAIPRPGIFVAVYAAWIKHII 715
OY 721 LTYKVPOS 728
Db 716 LTYKVPOS 723

RESULT 4

US-08-290-937B-1
; Sequence 1, Application US/08290937B
; Patent No. 5648233
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, KIYOJI
; APPLICANT: SHIMA, NOBUYUKI
; APPLICANT: MURAKAMI, AKIHIKO
; APPLICANT: GOTO, MASAOKI

APPLICANT: TSUDA, EISUKE
APPLICANT: MASUNAGA, HIROAKI
APPLICANT: TAKAHIRA, REIKO
APPLICANT: OOGAKI, FUMIKO
APPLICANT: UEDA, MASATSUGU
APPLICANT: HIGASHIO, KANJI
TITLE OF INVENTION: MODIFIED TCF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault
STREET: 125 High St.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,937B
FILING DATE: 19-AUG-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A.
REGISTRATION NUMBER: 32,503
REFERENCE/DOCKET NUMBER: FJN-022
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-937B-1

Query Match 99.1%; Score 4088.5; DB 1; Length 723;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 MMVTKLLPALLLOVHLHLPLTAIPYAEQORRRNTIHEFKSAKTLIKIDPALKIK 60
DB 1 MMVTKLLPALLLOVHLHLPLTAIPYAEQORRRNTIHEFKSAKTLIKIDPALKIK 60
61 TRKNTADOCANRCTRNKGLPFTCKAFVDFKARKOCLMFPNSMSSGVKKKEFGHEFDLYE 120
DB 61 TRKNTADOCANRCTRNKGLPFTCKAFVDFKARKOCLMFPNSMSSGVKKKEFGHEFDLYE 120
121 NKDYIRNCIIGKRSYKTSVITSGIKQCPWSSKIPHEHSFLPSSYKGLQENYCNP 180
DB 121 NKDYIRNCIIGKRSYKTSVITSGIKQCPWSSKIPHEHSFLPSSYKGLQENYCNP 180
121 NKDYIRNCIIGKRSYKTSVITSGIKQCPWSSKIPHEH-----SYGKDLQENYCNP 175
181 RGEESGPMCFSTNPEVREVCIDIPQCSVEECMTNGESYRGLMDHTEGSKICQRMHDTP 240
DB 181 RGEESGPMCFSTNPEVREVCIDIPQCSVEECMTNGESYRGLMDHTEGSKICQRMHDTP 240
176 RGEESGPMCFSTNPEVREVCIDIPQCSVEECMTNGESYRGLMDHTEGSKICQRMHDTP 235
241 HRHFLPERYPDKGFDNDYCNRPDQPRPWCYTLDPHTRMEYCAIKTCADNTMNDTVPL 300
DB 241 HRHFLPERYPDKGFDNDYCNRPDQPRPWCYTLDPHTRMEYCAIKTCADNTMNDTVPL 300
236 HRHFLPERYPDKGFDNDYCNRPDQPRPWCYTLDPHTRMEYCAIKTCADNTMNDTVPL 295
301 ETTGCIQGGGEGYRGTVNTINNGIIPCQWDSQYRPHEDMTPENFCKDLRENYCNPDS 360
DB 301 ETTGCIQGGGEGYRGTVNTINNGIIPCQWDSQYRPHEDMTPENFCKDLRENYCNPDS 360
296 ETTGCIQGGGEGYRGTVNTINNGIIPCQWDSQYRPHEDMTPENFCKDLRENYCNPDS 355
361 ESPPCFTDPNIRYGYGQINPCMSHGDQYRGNGKRYMGNLSQTRSGLTCSMDKME 420
DB 361 ESPPCFTDPNIRYGYGQINPCMSHGDQYRGNGKRYMGNLSQTRSGLTCSMDKME 420
356 ESPPCFTDPNIRYGYGQINPCMSHGDQYRGNGKRYMGNLSQTRSGLTCSMDKME 415

QY 421 DLHRIIFWEPDASKLENYCNRPDDDAHGPCYTGNPLIPMDYCPISRCEDDTPTIYNL 480
DB 416 DLHRIIFWEPDASKLENYCNRPDDDAHGPCYTGNPLIPMDYCPISRCEDDTPTIYNL 475
QY 481 DHPVISCARTKOLRVANGIPTRTNIGMWSLYRNKHIIGSSLIKESWYLTAROCFSPSRD 540
DB 476 DHPVISCARTKOLRVANGIPTRTNIGMWSLYRNKHIIGSSLIKESWYLTAROCFSPSRD 535
QY 541 LKDYEWAMGIDHVGGRGDECKOYLVNSQLYGPGSDLYIMKLARPAVLDDFVSTIDLP 600
DB 536 LKDYEWAMGIDHVGGRGDECKOYLVNSQLYGPGSDLYIMKLARPAVLDDFVSTIDLP 595
QY 601 NYGCTIPEKTSQSVYGMGYTGLINVDGLRYVAHLTYMGNEKCSOHRGKVTLNESEICAG 660
DB 596 NYGCTIPEKTSQSVYGMGYTGLINVDGLRYVAHLTYMGNEKCSOHRGKVTLNESEICAG 655
QY 661 AEKISGPECEDDYGGPLVCEQHKMRMVLGVYPGRCALIPNRPGLFVRAVAYAKIHKII 720
DB 656 AEKISGPECEDDYGGPLVCEQHKMRMVLGVYPGRCALIPNRPGLFVRAVAYAKIHKII 715
QY 721 LTKVYPOS 728
DB 716 LTKVYPOS 723

RESULT 5
US-08-404-643-1
Sequence 1, Application US/08404643
Patent No. 5658742
GENERAL INFORMATION:
APPLICANT: HIGASHIO, KANJI
APPLICANT: SHIMA, NOBUYUKI
APPLICANT: OOGAKI, FUMIKO
TITLE OF INVENTION: MONOCLONAL ANTIBODY
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
STREET: 53 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/404,643
FILING DATE: 15-MAR-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A.
REGISTRATION NUMBER: 32,503
REFERENCE/DOCKET NUMBER: FJN-034 (3999/35)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)-248-7100
TELEFAX: (617)-248-7100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-404-643-1

Query Match 99.1%; Score 4088.5; DB 1; Length 723;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 MWTLLPALLQVHLHLPLPAIPYAEQRRRNTIHEFKSATTLLIKIDPALKIK 60
1 MWTLLPALLQVHLHLPLPAIPYAEQRRRNTIHEFKSATTLLIKIDPALKIK 60
Db 1 MWTLLPALLQVHLHLPLPAIPYAEQRRRNTIHEFKSATTLLIKIDPALKIK 60
QY 61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPPNSMSGYKKEGHEFDLYE 120
61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPPNSMSGYKKEGHEFDLYE 120
Db 61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPPNSMSGYKKEGHEFDLYE 120
QY 121 NKDYIRNCIIIGKGRSYKGTVSITKSGIKQCPWSSMIPIHEHFLSSYRGKDLQENYCNRP 180
121 NKDYIRNCIIIGKGRSYKGTVSITKSGIKQCPWSSMIPIHEHFLSSYRGKDLQENYCNRP 180
Db 121 NKDYIRNCIIIGKGRSYKGTVSITKSGIKQCPWSSMIPIHEHFLSSYRGKDLQENYCNRP 180
QY 181 RGEEGGPMCTSNPEVREYVCDIPQCESEVECMTCNGESYRGKDLMTHTESGKICQRMHDOTP 240
181 RGEEGGPMCTSNPEVREYVCDIPQCESEVECMTCNGESYRGKDLMTHTESGKICQRMHDOTP 240
Db 181 RGEEGGPMCTSNPEVREYVCDIPQCESEVECMTCNGESYRGKDLMTHTESGKICQRMHDOTP 240
QY 241 HRHFLPERYPDKGFDDNYCNPDPDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 300
241 HRHFLPERYPDKGFDDNYCNPDPDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 300
Db 241 HRHFLPERYPDKGFDDNYCNPDPDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 300
QY 301 ETTECIGOGGEGYRGYVNTIWNIGIPCOQRMDSQYRPHEDMTPEHFKCDLRENYCNPDS 360
301 ETTECIGOGGEGYRGYVNTIWNIGIPCOQRMDSQYRPHEDMTPEHFKCDLRENYCNPDS 360
Db 301 ETTECIGOGGEGYRGYVNTIWNIGIPCOQRMDSQYRPHEDMTPEHFKCDLRENYCNPDS 360
QY 361 ESPMCTFTDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNSLQTSRGLTCSMMDKME 420
361 ESPMCTFTDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNSLQTSRGLTCSMMDKME 420
Db 361 ESPMCTFTDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNSLQTSRGLTCSMMDKME 420
QY 421 DLHRIIEMEPDASKLNEYCNPDDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 480
421 DLHRIIEMEPDASKLNEYCNPDDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 480
Db 421 DLHRIIEMEPDASKLNEYCNPDDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 480
QY 481 DHPVISCACKTQOLRVVNGIPIRTNIGMWVSLRYRNKHIICGSLIKESVWLTAROCFSPRD 540
481 DHPVISCACKTQOLRVVNGIPIRTNIGMWVSLRYRNKHIICGSLIKESVWLTAROCFSPRD 540
Db 481 DHPVISCACKTQOLRVVNGIPIRTNIGMWVSLRYRNKHIICGSLIKESVWLTAROCFSPRD 540
QY 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600
541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600
Db 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600
QY 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSOHHRGKVTLNSEICAG 660
601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSOHHRGKVTLNSEICAG 660
Db 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSOHHRGKVTLNSEICAG 660
QY 661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGRCALPVRPGIFVAVAYAAWIHKII 720
661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGRCALPVRPGIFVAVAYAAWIHKII 720
Db 661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGRCALPVRPGIFVAVAYAAWIHKII 720
QY 721 LTYKVPQS 728
721 LTYKVPQS 728
Db 721 LTYKVPQS 728
Db 716 LTYKVPQS 723

RESULT 6
US-09-194-326-1
; Sequence 1, Application US/09194326
; Patent No. 6306827
; GENERAL INFORMATION:
; APPLICANT: Kinoshita, Masahiko
; APPLICANT: Ogawa, Hitomi
; APPLICANT: Masanaga, Hiroaki
; APPLICANT: Kobayashi, Fumie
; APPLICANT: Yamaguchi, Kyoji
; APPLICANT: Higashio, Kanji
; TITLE OF INVENTION: Agent for Preventing and/or Treating Renal Disease
; FILE REFERENCE: FPN-069
; CURRENT APPLICATION NUMBER: US/09/194,326
; EARLIER APPLICATION NUMBER: PCT/JP98/01221
; EARLIER FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: JP 94989
; EARLIER FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1
; LENGTH: 723
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: TCF-II
US-09-194-326-1

Query Match 99.1%; Score 4088.5; DB 4; Length 723;
Best Local Similarity 99.38; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 MWTLLPALLQVHLHLPLPAIPYAEQRRRNTIHEFKSATTLLIKIDPALKIK 60
1 MWTLLPALLQVHLHLPLPAIPYAEQRRRNTIHEFKSATTLLIKIDPALKIK 60
Db 1 MWTLLPALLQVHLHLPLPAIPYAEQRRRNTIHEFKSATTLLIKIDPALKIK 60
QY 61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPPNSMSGYKKEGHEFDLYE 120
61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPPNSMSGYKKEGHEFDLYE 120
Db 61 TKKVTADQCANRCTRNKGLPFTCKAFVFDKARKOCLMFPPNSMSGYKKEGHEFDLYE 120
QY 121 NKDYIRNCIIIGKGRSYKGTVSITKSGIKQCPWSSMIPIHEHFLSSYRGKDLQENYCNRP 180
121 NKDYIRNCIIIGKGRSYKGTVSITKSGIKQCPWSSMIPIHEHFLSSYRGKDLQENYCNRP 180
Db 121 NKDYIRNCIIIGKGRSYKGTVSITKSGIKQCPWSSMIPIHEHFLSSYRGKDLQENYCNRP 180
QY 181 RGEEGGPMCTSNPEVREYVCDIPQCESEVECMTCNGESYRGKDLMTHTESGKICQRMHDOTP 240
181 RGEEGGPMCTSNPEVREYVCDIPQCESEVECMTCNGESYRGKDLMTHTESGKICQRMHDOTP 240
Db 181 RGEEGGPMCTSNPEVREYVCDIPQCESEVECMTCNGESYRGKDLMTHTESGKICQRMHDOTP 240
QY 241 HRHFLPERYPDKGFDDNYCNPDPDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 300
241 HRHFLPERYPDKGFDDNYCNPDPDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 300
Db 241 HRHFLPERYPDKGFDDNYCNPDPDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 300
QY 301 ETTECIGOGGEGYRGYVNTIWNIGIPCOQRMDSQYRPHEDMTPEHFKCDLRENYCNPDS 360
301 ETTECIGOGGEGYRGYVNTIWNIGIPCOQRMDSQYRPHEDMTPEHFKCDLRENYCNPDS 360
Db 301 ETTECIGOGGEGYRGYVNTIWNIGIPCOQRMDSQYRPHEDMTPEHFKCDLRENYCNPDS 360
QY 361 ESPMCTFTDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNSLQTSRGLTCSMMDKME 420
361 ESPMCTFTDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNSLQTSRGLTCSMMDKME 420
Db 361 ESPMCTFTDPNIRVGYCSQIPNCDMSHGQDCYRGNGKNYMGNSLQTSRGLTCSMMDKME 420
QY 421 DLHRIIEMEPDASKLNEYCNPDDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 480
421 DLHRIIEMEPDASKLNEYCNPDDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 480
Db 421 DLHRIIEMEPDASKLNEYCNPDDAGHPCWCTYTLDPHTRWECYCAIKTCADNTMNDTVP 480
QY 481 DHPVISCACKTQOLRVVNGIPIRTNIGMWVSLRYRNKHIICGSLIKESVWLTAROCFSPRD 540
481 DHPVISCACKTQOLRVVNGIPIRTNIGMWVSLRYRNKHIICGSLIKESVWLTAROCFSPRD 540
Db 481 DHPVISCACKTQOLRVVNGIPIRTNIGMWVSLRYRNKHIICGSLIKESVWLTAROCFSPRD 540
QY 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600
541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600
Db 541 LKDYEAWLGIHDVHGRDEKCKOVLNVSQLYGPBGSDVLMKLARPAVLDDPFTSTIDLP 600
QY 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSOHHRGKVTLNSEICAG 660
601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSOHHRGKVTLNSEICAG 660
Db 601 NYGCTIPEKTSQSYGMYGTGLINVDGLLRVAHLIYIMNEKCSOHHRGKVTLNSEICAG 660
QY 661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGRCALPVRPGIFVAVAYAAWIHKII 720
661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGRCALPVRPGIFVAVAYAAWIHKII 720
Db 661 AEKIGSGCEBDYGGPLVCEQHKRMVGLVIVPGRCALPVRPGIFVAVAYAAWIHKII 720
QY 721 LTYKVPQS 728
721 LTYKVPQS 728
Db 721 LTYKVPQS 728
Db 716 LTYKVPQS 723

RESULT 7
US-08-290-937B-2
; Sequence 2, Application US/08290937B
; Patent No. 5648233
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, KYOJI

```

? APPLICANT: SHIMA, NOBUYUKI
? APPLICANT: MURAKAMI, AKIHIKO
? APPLICANT: GOTO, MASAOKI
? APPLICANT: TSUDA, EISUKE
? APPLICANT: MASUNAGA, HIROAKI
? APPLICANT: TAKAHIRA, REIKO
? APPLICANT: OOGAKI, FUMIKO
? APPLICANT: UEDA, MASATOSUGU
? APPLICANT: HIGASHIO, KANJI
? TITLE OF INVENTION: MODIFIED TCF
? NUMBER OF SEQUENCES: 13
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Testa, Hurwitz & Thibault
? STREET: 125 High St.
? CITY: Boston
? STATE: MA
? COUNTRY: USA
? ZIP: 02110
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/290,937B
? FILING DATE: 19-AUG-1994
? CLASSIFICATION: 530
? ATTORNEY/AGENT INFORMATION:
? NAME: CAMPBELL, PAULA A.
? REGISTRATION NUMBER: 32,503
? REFERENCE/DOCKET NUMBER: FJN-022
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 248-7000
? TELEFAX: (617) 248-7100
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 723 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-290-937B-2
?
Query Match 99.08; Score 4085.5; DB 1; Length 723;
Best Local Similarity 99.28; Pred. No. 0;
Matches 722; Conservative 1; Mismatches 0; Indels 5; Gaps 1;
1 MMVTKRLPALLOHVLHLPLPIPAEGORRRNTIHEFKSAKTTLIKIDPALKIK 60
1 MMVTKRLPALLOHVLHLPLPIPAEGORRRNTIHEFKSAKTTLIKIDPALKIK 60
1 MMVTKRLPALLOHVLHLPLPIPAEGORRRNTIHEFKSAKTTLIKIDPALKIK 60
?
QY 61 TRKYNTADOCANRCTRNKGLPFTCKAFVFDKARKQCLMFPTNSMSSGVKKEGHEFDLYE 120
?
DB 61 TRKYNTADOCANRCTRNKGLPFTCKAFVFDKARKQCLMFPTNSMSSGVKKEGHEFDLYE 120
?
QY 121 NKDYTRNCIIGKRSYKSTVSTTKSGIKCQPMSSMTIPEHSHLPSSYSGKDLQENYCNRP 180
?
DB 121 NKDYTRNCIIGKRSYKSTVSTTKSGIKCQPMSSMTIPEHSHLPSSYSGKDLQENYCNRP 180
?
DB 121 NKDYTRNCIIGKRSYKSTVSTTKSGIKCQPMSSMTIPEHSHLPSSYSGKDLQENYCNRP 175
?
QY 181 RGEEGSPCFTSNPEVREVCODIPCCSEFECMTNGESYRGIMLDTESKICQORDHOTP 240
?
DB 181 RGEEGSPCFTSNPEVREVCODIPCCSEFECMTNGESYRGIMLDTESKICQORDHOTP 240
?
DB 176 RGEEGSPCFTSNPEVREVCODIPCCSEFECMTNGESYRGIMLDTESKICQORDHOTP 235
?
QY 241 HRHKFLPERYPDKGFDNDYCRNPDGQPRPWCYTLDPHTRMEYCAIKTCADNTMNDTVDPL 300
?
DB 236 HRHKFLPERYPDKGFDNDYCRNPDGQPRPWCYTLDPHTRMEYCAIKTCADNTMNDTVDPL 295
?
QY 301 ETTBETIOGEGESYRGTVNTIMNGICQORDSQYRPHEDMTPENFCKDKDLRENYCNPDS 360
?
DB 296 ETTBETIOGEGESYRGTVNTIMNGICQORDSQYRPHEDMTPENFCKDKDLRENYCNPDS 355
?
QY 361 ESPWCFITTDPNIRVGCISQIPNCDSHGDQYRGNGKNYMGILSGTSGLTCSMDKNME 420

```

```

DB 356 ESPWCFITTDPNIRVGCISQIPNCDSHGDQYRGNGKNYMGILSGTSGLTCSMDKNME 415
?
QY 421 DLHRHIFWEPDASKLNNENYCRNPDGDAHGFWCYTGNPLIPWDYCPISRCBGDTTPITVNI 480
?
DB 416 DLHRHIFWEPDASKLNNENYCRNPDGDAHGFWCYTGNPLIPWDYCPISRCBGDTTPITVNI 475
?
QY 481 DHPVISCARKTOLRVYNGITPTNIGMVSRLRYRNKHICGSLIKESWYLTARQCPSPRD 540
?
DB 476 DHPVISCARKTOLRVYNGITPTNIGMVSRLRYRNKHICGSLIKESWYLTARQCPSPRD 535
?
QY 541 LKDYEAMLGIDHYHGRGDEKQVLANYSQLYVGPESDLYLMKLARPAVLDDFVSTIDLP 600
?
DB 536 LKDYEAMLGIDHYHGRGDEKQVLANYSQLYVGPESDLYLMKLARPAVLDDFVSTIDLP 595
?
QY 601 NYGCTIPEKTSYGVGMYGTYGLINYGDLRVAAHLYIMGANKCSQHRGKTYLWSEFCAG 660
?
DB 596 NYGCTIPEKTSYGVGMYGTYGLINYGDLRVAAHLYIMGANKCSQHRGKTYLWSEFCAG 655
?
QY 661 AEKIGSGPCGGDYGGLVPCDQHKRMVLYGVIPRGCAIPNRGIEFVAVYAKWIKIT 720
?
DB 656 AEKIGSGPCGGDYGGLVPCDQHKRMVLYGVIPRGCAIPNRGIEFVAVYAKWIKIT 715
?
QY 721 LTYKVPQS 728
?
DB 716 LTYKVPQS 723
?

```

3

```

RESULT 8
US-08-290-937B-3
: Sequence 3, Application US/08290937B
: Patent No. 5648233
: GENERAL INFORMATION:
: APPLICANT: YAMAGUCHI, KYOJI
: APPLICANT: SHIMA, NOBUYUKI
: APPLICANT: MURAKAMI, AKIHIKO
: APPLICANT: GOTO, MASAOKI
: APPLICANT: TSUDA, EISUKE
: APPLICANT: MASUNAGA, HIROAKI
: APPLICANT: TAKAHIRA, REIKO
: APPLICANT: OOGAKI, FUMIKO
: APPLICANT: UEDA, MASATOSUGU
: APPLICANT: HIGASHIO, KANJI
: TITLE OF INVENTION: MODIFIED TCF
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Testa, Hurwitz & Thibault
: STREET: 125 High St.
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/290,937B
: FILING DATE: 19-AUG-1994
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: CAMPBELL, PAULA A.
: REGISTRATION NUMBER: 32,503
: REFERENCE/DOCKET NUMBER: FJN-022
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 248-7000
: TELEFAX: (617) 248-7100
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 723 amino acids
: TYPE: amino acid
: STRANDEDNESS: single

```

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-290-937B-3

Query Match 98.9%; Score 4079.5; DB 1; Length 723;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 721; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

```

QY 1 MWYTKLLPALLQVLLHLLHLLPLAIPYAGQRRKRNTHHEFKSATTTLIKIDPAKIK 60
   |||||||
Db 1 MWYTKLLPALLQVLLHLLHLLPLAIPYAGQRRKRNTHHEFKSATTTLIKIDPAKIK 60
QY 61 TKKYNADQCANRCTRNKGLPFTCKAFVFDKARKQCLMFPPNSMSGYKKEFGHEFDLYE 120
   |||||||
Db 61 TKKYNADQCANRCTRNKGLPFTCKAFVFDKARKQCLMFPPNSMSGYKKEFGHEFDLYE 120
QY 121 NKDIYRNCIIIGKGRSYGTAVSITKSGIKCOPWSSMIPHEHSFLPSSYRGDLOENYCRNP 180
   |||||||
Db 121 NKDIYRNCIIIGKGRSYGTAVSITKSGIKCOPWSSMIPHEH-----SYRKDLOENYCRNP 175
QY 181 RGEGBGPMCTSNPEVREYECDDIPQCESEVCEMTCNGESYRGLMDHTSGKICQRMWDQTP 240
   |||||||
Db 176 RGEGBGPMCTSNPEVREYECDDIPQCESEVCEMTCNGESYRGLMDHTSGKICQRMWDQTP 235
QY 241 HRHKFLPERYPDKGFDNRYCNPDPGQPRPWCYTLDPTRMWEYCAIKTCADNTMDTVPPL 300
   |||||||
Db 236 HRHKFLPERYPDKGFDNRYCNPDPGQPRPWCYTLDPTRMWEYCAIKTCADNTMDTVPPL 295
QY 301 ETTECIOGQGBGYRGVNTWNGIIPCQRMDSQYRHEHDMTPENFKCKDLRENYCRNPDGS 360
   |||||||
Db 296 ETTECIOGQGBGYRGVNTWNGIIPCQRMDSQYRHEHDMTPENFKCKDLRENYCRNPDGS 355
QY 361 ESPMCFETTDPIRIVGYCSQIIPNCDSHGOCDYRGNGKNGMNLSTQTSGLTCSMWDKME 420
   |||||||
Db 356 ESPMCFETTDPIRIVGYCSQIIPNCDSHGOCDYRGNGKNGMNLSTQTSGLTCSMWDKME 415
QY 421 DLHRHIFWEPDASKLNNRYCNPDDDAHGPWCYTGNPLIPDYCPISRCEDDTPITYNL 480
   |||||||
Db 416 DLHRHIFWEPDASKLNNRYCNPDDDAHGPWCYTGNPLIPDYCPISRCEDDTPITYNL 475
QY 481 DHPVISCARKTKQLRVNGIPLRTNIGMWSLRYRNKHIICGSLIKESWVLTAROCFSPSRD 540
   |||||||
Db 476 DHPVISCARKTKQLRVNGIPLRTNIGMWSLRYRNKHIICGSLIKESWVLTAROCFSPSRD 535
QY 541 LKDYEAWLGIHDVHGRDECKQVLYNLSQVLYGEGSDVLYMLKARAVLDDFVSTIDLP 600
   |||||||
Db 536 LKDYEAWLGIHDVHGRDECKQVLYNLSQVLYGEGSDVLYMLKARAVLDDFVSTIDLP 595
QY 601 NYGCTIPEKTSASYGMYGTGLINYDGLLRVAHLXYIMGNEKCSOHHHGKVTLLNESEICAG 660
   |||||||
Db 596 NYGCTIPEKTSASYGMYGTGLINYDGLLRVAHLXYIMGNEKCSOHHHGKVTLLNESEICAG 655
QY 661 AEKIGSPCEGDYGGPLVCEQHKMRYLVGIIVPGGCAIPNRPGIFVRAVAYAKIHKII 720
   |||||||
Db 656 AEKIGSPCEGDYGGPLVCEQHKMRYLVGIIVPGGCAIPNRPGIFVRAVAYAKIHKII 715
QY 721 LTYKVPQS 728
   |||||||
Db 716 LTYKVPQS 723

```

RESULT 9
US-08-030-410-3
Sequence 3, Application US/08030410
Patent No. 6221359

GENERAL INFORMATION:
APPLICANT: Komiya, Atsushi
APPLICANT: Nakahata, Tatsutoshi
APPLICANT: Kubo, Tetsuo
APPLICANT: Tanaka, Ryunei
APPLICANT: Kawano, Genji
APPLICANT: Sudo, Tetsuo

```

APPLICANT: Sano, Emiko
APPLICANT: Kojima, Katsuki
TITLE OF INVENTION: HEMOPHETIC STEM CELL MULTIPLIER
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/030,410
FILING DATE: 19930521
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8898
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 723 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-030-410-3

```

Query Match 98.8%; Score 4077.5; DB 4; Length 723;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 720; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

```

QY 1 MWYTKLLPALLQVLLHLLHLLPLAIPYAGQRRKRNTHHEFKSATTTLIKIDPAKIK 60
   |||||||
Db 1 MWYTKLLPALLQVLLHLLHLLPLAIPYAGQRRKRNTHHEFKSATTTLIKIDPAKIK 60
QY 61 TKKYNADQCANRCTRNKGLPFTCKAFVFDKARKQCLMFPPNSMSGYKKEFGHEFDLYE 120
   |||||||
Db 61 TKKYNADQCANRCTRNKGLPFTCKAFVFDKARKQCLMFPPNSMSGYKKEFGHEFDLYE 120
QY 121 NKDIYRNCIIIGKGRSYGTAVSITKSGIKCOPWSSMIPHEHSFLPSSYRGDLOENYCRNP 180
   |||||||
Db 121 NKDIYRNCIIIGKGRSYGTAVSITKSGIKCOPWSSMIPHEH-----SYRKDLOENYCRNP 175
QY 181 RGEGBGPMCTSNPEVREYECDDIPQCESEVCEMTCNGESYRGLMDHTSGKICQRMWDQTP 240
   |||||||
Db 176 RGEGBGPMCTSNPEVREYECDDIPQCESEVCEMTCNGESYRGLMDHTSGKICQRMWDQTP 235
QY 241 HRHKFLPERYPDKGFDNRYCNPDPGQPRPWCYTLDPTRMWEYCAIKTCADNTMDTVPPL 300
   |||||||
Db 236 HRHKFLPERYPDKGFDNRYCNPDPGQPRPWCYTLDPTRMWEYCAIKTCADNTMDTVPPL 295
QY 301 ETTECIOGQGBGYRGVNTWNGIIPCQRMDSQYRHEHDMTPENFKCKDLRENYCRNPDGS 360
   |||||||
Db 296 ETTECIOGQGBGYRGVNTWNGIIPCQRMDSQYRHEHDMTPENFKCKDLRENYCRNPDGS 355
QY 361 ESPMCFETTDPIRIVGYCSQIIPNCDSHGOCDYRGNGKNGMNLSTQTSGLTCSMWDKME 420
   |||||||
Db 356 ESPMCFETTDPIRIVGYCSQIIPNCDSHGOCDYRGNGKNGMNLSTQTSGLTCSMWDKME 415
QY 421 DLHRHIFWEPDASKLNNRYCNPDDDAHGPWCYTGNPLIPDYCPISRCEDDTPITYNL 480
   |||||||
Db 416 DLHRHIFWEPDASKLNNRYCNPDDDAHGPWCYTGNPLIPDYCPISRCEDDTPITYNL 475

```



```

Db 1 MWTKLLPALLQHVLLHLLPLAIPYAEQORRRNTIHEFKSAKTLIKIDPALAIA 60
QY 61 TKKVTADOCANRCTRNKGLPFTCKAFVFDKARKOCLWFPPNSMSGYKKEGHEFDLYE 120
Db 61 TAAVNTADOCANRCTRNKGLPFTCKAFVFDKARKOCLWFPPNSMSGYKKEGHEFDLYE 120
QY 121 NKDYIRNCTICKGRSYKTSVITNSGKICQPMSSMIPEHNSFLPSSYRGKDLQENYCRNP 180
Db 121 NKDYIRNCTICKGRSYKTSVITNSGKICQPMSSMIPEHNSFLPSSYRGKDLQENYCRNP 175
QY 181 RGEESGPMCFSTSNPEVREVDIPQCSFEVCEMTGCESYRGMLMHTESGKICQRMDHOTP 240
Db 176 RGEESGPMCFSTSNPEVREVDIPQCSFEVCEMTGCESYRGMLMHTESGKICQRMDHOTP 235
QY 241 HRHFLPERYDCKGFDNDYCRNPDGQPRPWCYTLDPHTRMEYCAIKTKADTMNDTDPVL 300
Db 236 HRHFLPERYDCKGFDNDYCRNPDGQPRPWCYTLDPHTRMEYCAIKTKADTMNDTDPVL 295
QY 301 ETTECIGQGGEGYRGTVNTIWNIGIPCCQRMDSQYPHEHDMTPENFKCKDLRENYCRNPDGS 360
Db 296 ETTECIGQGGEGYRGTVNTIWNIGIPCCQRMDSQYPHEHDMTPENFKCKDLRENYCRNPDGS 355
QY 361 ESPMCFETDPMIRVGYCSQIIPNCMSHGQDCYRGNGKNYMNLSQTRSGGLTCSMDKXME 420
Db 356 ESPMCFETDPMIRVGYCSQIIPNCMSHGQDCYRGNGKNYMNLSQTRSGGLTCSMDKXME 415
QY 421 DLHRIEWEPPASKLNEYCRNPDGDAHGPCYTGNNPLIPMDYCIISRCEGDTPTIYNL 480
Db 416 DLHRIEWEPPASKLNEYCRNPDGDAHGPCYTGNNPLIPMDYCIISRCEGDTPTIYNL 475
QY 481 DHPVISCATKQLRVYNGIPIRTNIGMNVSLRYRNKHAICGSSLIKESWVLTARQCFPSRD 540
Db 476 DHPVISCATKQLRVYNGIPIRTNIGMNVSLRYRNKHAICGSSLIKESWVLTARQCFPSRD 535
QY 541 LKQDEAWLGIHDVHGRGDEKCKQVLTANVQLYYGPBGSDVLYMLKARPAVLDDFVSTIDLP 600
Db 536 LKQDEAWLGIHDVHGRGDEKCKQVLTANVQLYYGPBGSDVLYMLKARPAVLDDFVSTIDLP 595
QY 601 NYGCTIPEKTSQSYGMYGTGLINVDGLRVANLXIMGENKCSQHHRRKVTLLNESEICAG 660
Db 596 NYGCTIPEKTSQSYGMYGTGLINVDGLRVANLXIMGENKCSQHHRRKVTLLNESEICAG 655
QY 661 AEKIGSGCEBDYGGPLVCEQHKMVMVLGVTPRGCAIIPNRPGLFVAVAYAKMIHKII 720
Db 656 AEKIGSGCEBDYGGPLVCEQHKMVMVLGVTPRGCAIIPNRPGLFVAVAYAKMIHKII 715
QY 721 LTYKVPQS 728
Db 716 LTYKVPQS 723

```

```

RESULT 12
: Sequence 50, Application US/08460890A
: Patent No. 5994109
: GENERAL INFORMATION:
: APPLICANT: Moo, Savio L.C.
: APPLICANT: Smith, Louis C.
: APPLICANT: Cristiano, Richard J.
: APPLICANT: Gotcheak, Stephen
: TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
: NUMBER OF SEQUENCES: 65
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: CITY: Suite 4700
: STATE: Los Angeles
: COUNTRY: California
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

```

```

: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: FastSeq for Windows 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/460,890A
: FILING DATE: June 5, 1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/167,641
: FILING DATE: December 14, 1993
: APPLICATION NUMBER: 07/855,389
: FILING DATE: March 20, 1992
: APPLICATION NUMBER: PCT/US93/02725
: FILING DATE: March 19, 1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Walburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 212/066
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 50:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 697 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-460-890A-50

Query Match 96.2%; Score 3968; DB 2; Length 697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 572 YGPEGSDLVLMKLARPAVLDFVSTIDLPNGCTTPEKTSQSVYGMVYTGILINDGLLRV 631
DB 541 YGPEGSDLVLMKLARPAVLDFVSTIDLPNGCTTPEKTSQSVYGMVYTGILINDGLLRV 600
QY 632 AHLTYMGNEKCSQHHRGVTTLNSESICAGAEKIGSGPEGDYGGLVCEQHMRVNLGYI 691
DB 601 AHLTYMGNEKCSQHHRGVTTLNSESICAGAEKIGSGPEGDYGGLVCEQHMRVNLGYI 660
QY 692 YGRCGCAIPNRPGRIFVRAVAYAKWIKIILTYKVPQS 728
DB 661 YGRCGCAIPNRPGRIFVRAVAYAKWIKIILTYKVPQS 697

RESULT 13
US-08-167-641C-50
; Sequence 50, Application US/08167641C
; Patent No. 6033884
; GENERAL INFORMATION:
; APPLICANT: WOO, Saviio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristliano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08167,641C
; FILING DATE: December 14, 1993
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 07/855,389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 205/012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 697 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-167-641C-50

Query Match 96.2%; Score 3968; DB 3; Length 697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 QRRRRNTTIEFKKSAKTTLIKIDPALKITKKNVTADQCANCRTRKNGCLPFTCKAVFVK 91
DB 1 QRRRRNTTIEFKKSAKTTLIKIDPALKITKKNVTADQCANCRTRKNGCLPFTCKAVFVK 60

QY 92 ARKOCLEPFENSMSSGVAKKEGHEFDLYENKOTYRNCILGKSGYKTSITKSGIKCP 151
DB 61 ARKOCLEPFENSMSSGVAKKEGHEFDLYENKOTYRNCILGKSGYKTSITKSGIKCP 120
QY 152 WSMMPHEHSFLPSSYRKCDLOENYCRNPGEEGPWCFTSNPEVREYCDIPCCSEVEC 211
DB 121 WSMMPHEHSFLPSSYRKCDLOENYCRNPGEEGPWCFTSNPEVREYCDIPCCSEVEC 180
QY 212 MTCNGESYRGLMDHTESGKICQRMWDQTPHRRKFLPERYPDKGFDNYCRNPDGQPRPC 271
DB 181 MTCNGESYRGLMDHTESGKICQRMWDQTPHRRKFLPERYPDKGFDNYCRNPDGQPRPC 240
QY 272 YTLDPHTWMEYCAIKTCADNTMNDTVPLETTTECIGQGBEGYKGTVNTIMNGIFCQMRDS 331
DB 241 YTLDPHTWMEYCAIKTCADNTMNDTVPLETTTECIGQGBEGYKGTVNTIMNGIFCQMRDS 300
QY 332 QYHEHDMTPENFKCKDIRENYCRNPDESPPWCFTTDPNIRVYCSQIPNCDMSHGQDC 391
DB 301 QYHEHDMTPENFKCKDIRENYCRNPDESPPWCFTTDPNIRVYCSQIPNCDMSHGQDC 360
QY 392 YRGNKNYMGNLSTQSGLTCSMDKNMEDLHRHIFWEPDASKINENYCRNPDDDAHGPW 451
DB 361 YRGNKNYMGNLSTQSGLTCSMDKNMEDLHRHIFWEPDASKINENYCRNPDDDAHGPW 420
QY 452 CYTGNPILPMDYCPISRCGDTPTIVLNDHPVISCATKOLRVNGIPTRTNIGMAYSL 511
DB 421 CYTGNPILPMDYCPISRCGDTPTIVLNDHPVISCATKOLRVNGIPTRTNIGMAYSL 480
QY 512 RYRNKHICGSLIKESWVLTRACQPPSRDLKDYEAMLGIDHVGHGDECKQVNLVSOLV 571
DB 481 RYRNKHICGSLIKESWVLTRACQPPSRDLKDYEAMLGIDHVGHGDECKQVNLVSOLV 540
QY 572 YGPEGSDLVLMKLARPAVLDFVSTIDLPNGCTTPEKTSQSVYGMVYTGILINDGLLRV 631
DB 541 YGPEGSDLVLMKLARPAVLDFVSTIDLPNGCTTPEKTSQSVYGMVYTGILINDGLLRV 600
QY 632 AHLTYMGNEKCSQHHRGVTTLNSESICAGAEKIGSGPEGDYGGLVCEQHMRVNLGYI 691
DB 601 AHLTYMGNEKCSQHHRGVTTLNSESICAGAEKIGSGPEGDYGGLVCEQHMRVNLGYI 660
QY 692 YGRCGCAIPNRPGRIFVRAVAYAKWIKIILTYKVPQS 728
DB 661 YGRCGCAIPNRPGRIFVRAVAYAKWIKIILTYKVPQS 697

RESULT 14
US-08-460-971A-50
; Sequence 50, Application US/08460971A
; Patent No. 6150168
; GENERAL INFORMATION:
; APPLICANT: WOO, Saviio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristliano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/460,971A
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/167,641
FILING DATE: December 14, 1993
APPLICATION NUMBER: 07/855,389
FILING DATE: March 20, 1992
APPLICATION NUMBER: PCT/US93/02725
FILING DATE: March 19, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 212/063
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 697 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-460-971A-50

Query Match 96.2%; Score 3968; DB 4; Length 697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

32 ORRRNTIHEFKSATTLLIKIDPALIKITKKVNTADOCANRCKRNKGLPTCKAFEDK 91
1 ORRRNTIHEFKSATTLLIKIDPALIKITKKVNTADOCANRCKRNKGLPTCKAFEDK 60
92 ARKOCLEFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKTVSITKSGICOP 151
61 ARKOCLEFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKTVSITKSGICOP 120
152 WSMITHFHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVREYVCDIPQCSVEEC 211
121 WSMITHFHSFLPSSYRGKDLQENYCRNPRGEGGFWCTSNPEVREYVCDIPQCSVEEC 180
212 MTCNGESYGLMHTSGKICQWMDHQTPIRHRKFLPERYPDKGFDQYCNPPGQPPWC 271
181 MTCNGESYGLMHTSGKICQWMDHQTPIRHRKFLPERYPDKGFDQYCNPPGQPPWC 240
272 YLDPHTRMFYCAIKTCADNTMTDTPLETTTCIOGQGGYGTVNTINNGIIPCORWDS 331
241 YLDPHTRMFYCAIKTCADNTMTDTPLETTTCIOGQGGYGTVNTINNGIIPCORWDS 300
332 QYRHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCDMSHGQDC 391
301 QYRHEHDMTPENFKCKDLRENYCRNPDGSESPWCFTTDPNIRVGYCSQIIPNCDMSHGQDC 360
392 YRGNKNGNYMGNLSQTSRGLTCSMWDKMMEDLHRHITWEPASKLNEYCNPPDDDAHGPW 451
361 YRGNKNGNYMGNLSQTSRGLTCSMWDKMMEDLHRHITWEPASKLNEYCNPPDDDAHGPW 420
452 CYTGNPLIPMDYCPISRCEDGTPPTIVNDHPYISCAKTKQLRVNNGIIPRTNIGMWVSL 511
421 CYTGNPLIPMDYCPISRCEDGTPPTIVNDHPYISCAKTKQLRVNNGIIPRTNIGMWVSL 480
512 RYRNKHIICGSLIKESWVLTAROCFSPRDLKDYEAMLGIDHVGREGDECKQYLVANSQLV 571
481 RYRNKHIICGSLIKESWVLTAROCFSPRDLKDYEAMLGIDHVGREGDECKQYLVANSQLV 540
572 YRGGESDLYMKTARAVYLDDEFVSTIDLRYNGCTIPEKTSQSVYGMQYTGILLINVDGLRY 631
541 YRGGESDLYMKTARAVYLDDEFVSTIDLRYNGCTIPEKTSQSVYGMQYTGILLINVDGLRY 600
632 AHIYINGNEKCSQHHRGKVTLLNSESICAGAEKIGSGPCEDGYGGLVCEQHKKRMVGLVI 691

|||||
Db 601 AHIYINGNEKCSQHHRGKVTLLNSESICAGAEKIGSGPCEDGYGGLVCEQHKKRMVGLVI 660
QY 692 VPRGCAIPNRPGLFVAVYAKYKIHKILITRYVPOS 728
Db 661 VPRGCAIPNRPGLFVAVYAKYKIHKILITRYVPOS 697

RESULT 15

US-08-462-040-50
Sequence 50, Application US/08462040
Patent No. 6177554
GENERAL INFORMATION:
APPLICANT: WOO, Savio L.C.
APPLICANT: Smith, Louis C.
APPLICANT: Cristiano, Richard J.
TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
TITLE OF INVENTION: METHODS OF USE
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FASTSEQ for Windows 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,040

FILING DATE: June 5, 1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/167,641

FILING DATE: December 14, 1993

APPLICATION NUMBER: 07/855,389

FILING DATE: March 20, 1992

APPLICATION NUMBER: PCT/US93/02725

FILING DATE: March 19, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 212/078

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 697 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-462-040-50

Query Match 96.2%; Score 3968; DB 4; Length 697;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

32 ORRRNTIHEFKSATTLLIKIDPALIKITKKVNTADOCANRCKRNKGLPTCKAFEDK 91
1 ORRRNTIHEFKSATTLLIKIDPALIKITKKVNTADOCANRCKRNKGLPTCKAFEDK 60
92 ARKOCLEFPNSMSSGVKKEFGHEFDLYENKDYIRNCIIIGKRSYKTVSITKSGICOP 151

Db 61 ARKQCLMFFPNSNSGVKKEFGHEFDLYENKDYIRNCLIGKGRSYKGVSTKSGIKOP 120
QY 152 WSSMTIPEHESFLPSSYRGKDLQENYCRNPRGEGGPMCFTSNPEVRYEYCDIPQCSVEVC 211
Db 121 WSSMTIPEHESFLPSSYRGKDLQENYCRNPRGEGGPMCFTSNPEVRYEYCDIPQCSVEVC 180
QY 212 MTCNGESYRGMDHTESGKICQRMHDQTPHREKELPERYPDKGFDNNYCRNPDGQPRMVC 271
Db 181 MTCNGESYRGMDHTESGKICQRMHDQTPHREKELPERYPDKGFDNNYCRNPDGQPRMVC 240
QY 272 YTLDPHTRMWYCAIKTCADMTMNDTVPLETTTCIOGOGGYRGYNTVNTIMNGIIPCQRMDS 331
Db 241 YTLDPHTRMWYCAIKTCADMTMNDTVPLETTTCIOGOGGYRGYNTVNTIMNGIIPCQRMDS 300
QY 332 QYHEHDMTPENFKCKDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIIPNCDMSHGQDC 391
Db 301 QYHEHDMTPENFKCKDLRENYCRNPDGSESPMCFTTDPNIRVGYCSQIIPNCDMSHGQDC 360
QY 392 YRGNGKNYMGNLQOTRSGLTCSMMDKNMEDLHRHIFWEPDASKLNNYCRNPDGDAHGPW 451
Db 361 YRGNGKNYMGNLQOTRSGLTCSMMDKNMEDLHRHIFWEPDASKLNNYCRNPDGDAHGPW 420
QY 452 CYTGNPLIPMDYCPISRCEDGDTPTIYNLDHPYISCAKTKQLRVNNGIPTRTNIGMNVSL 511
Db 421 CYTGNPLIPMDYCPISRCEDGDTPTIYNLDHPYISCAKTKQLRVNNGIPTRTNIGMNVSL 480
QY 512 RYRNKHIQGSGLIKESVNLTAROCPSRDLKDYEAWLGIHDVHGRGDECKOYLVNSQV 571
Db 481 RYRNKHIQGSGLIKESVNLTAROCPSRDLKDYEAWLGIHDVHGRGDECKOYLVNSQV 540
QY 572 YGPGSDLVLMKLARPAVLDDFVSTIDLPNYGCITPEKTSQSVYGMGYTGLINYDGLRV 631
Db 541 YGPGSDLVLMKLARPAVLDDFVSTIDLPNYGCITPEKTSQSVYGMGYTGLINYDGLRV 600
QY 632 AHLIYMGNEKCSQHHRGKVTLNESEICAGAEKIGSPCEGDYGGPLYCEQHKMRMVLGYI 691
Db 601 AHLIYMGNEKCSQHHRGKVTLNESEICAGAEKIGSPCEGDYGGPLYCEQHKMRMVLGYI 660
QY 692 VPGRCALPNRPGTFVRVAYYAKWIKIILLYKVPQS 728
Db 661 VPGRCALPNRPGTFVRVAYYAKWIKIILLYKVPQS 697

Search completed: June 18, 2002, 17:47:21
Job time: 242 sec